

现代数学译丛 24

变分分析与广义微分

II：应用

〔美〕Boris S.Mordukhovich 著

李春 王炳武 赵亚莉 王东 译



科学出版社

现代数学译丛 24

变分分析与广义微分

II：应用

〔美〕 Boris S. Mordukhovich 著

李 春 王炳武 赵亚莉 王 东 译



科学出版社

北 京



图字：01-2013-8292 号

内 容 简 介

《变分分析与广义微分》是现代变分分析创始人之一的美国州立韦恩大学(Wayne State University) 的 Boris S. Mordukhovich 教授的最新专著，涵盖了无穷维空间中变分分析的最新成果及其应用. 原著分两卷，上卷阐述无穷维变分分析的基础理论，下卷则讨论在最优化、控制和经济学等各方面的应用. 第 5 章系统探讨了无穷维空间上的光滑和非光滑约束优化与均衡问题. 第 6 章和第 7 章论述了变分分析在动态最优化和最优控制上的应用. 其中第 6 章研究由常微分动力系统控制的最优控制问题；第 7 章讨论分布参数控制系统. 第 8 章提供了变分分析在福利经济学中的应用.

本书主要面向非线性分析、最优化、均衡、控制和对策论、泛函微分方程、数理经济等相关专业的高年级本科生和研究生，也可供运筹学、系统分析、力学、工程和经济学中涉及变分方法的研究人员和工程技术人员参考.

图书在版编目(CIP)数据

变分分析与广义微分 Ⅱ：应用/(美)莫尔杜霍维奇(Mordukhovich. B. S.) 著；李春等译. —北京：科学出版社，2014. 1
(现代数学译丛；24)
书名原文：Variational analysis and generalized differentiation Ⅱ: applications
ISBN 978-7-03-039263-3
Ⅰ. 变… Ⅱ. ① 莫… ② 李… Ⅲ. ① 泛函分析 Ⅳ. O177
中国版本图书馆 CIP 数据核字(2013) 第 290779 号

责任编辑：李 欣／责任校对：郑金红 刘亚琦
责任印制：赵德静／封面设计：王 浩

科学出版社 出版
北京东黄城根北街 16 号
邮政编码：100717
http://www.sciencep.com

文林印务有限公司 印刷
科学出版社发行 各地新华书店经销
*
2014 年 1 月第 一 版 开本：720 × 1000 1/16
2014 年 1 月第一次印刷 印张：34
字数：660 000

定价：138.00 元
(如有印装质量问题，我社负责调换)

译 者 序

《变分分析与广义微分》是现代变分分析创始人之一的美国州立韦恩大学 (Wayne State University) 的 Boris S. Mordukhovich 教授的最新专著, 涵盖了无穷维空间中变分分析的最新成果及其应用. 原著分两卷, 上卷阐述无穷维变分分析的基础理论, 下卷则讨论在最优化、控制和经济学等各方面的应用. 上卷中译本已于 2011 年 9 月出版, 这里翻译的是下卷.

本卷的翻译始于 2011 年夏天美国东密歇根大学 (Eastern Michigan University) 的王炳武教授 (上卷译者之一) 访问渤海大学期间. 当时正值本书上卷译讫, 渤海大学李春教授提议主持下卷的翻译, 并得到了王炳武教授和渤海大学赵亚莉教授 (上卷译者之一) 的支持. 很荣幸美国州立 Fayetteville 大学 (Fayetteville State University) 的王东教授也参与到翻译中来. 王炳武和王东都是 Mordukhovich 的学生, 也一直是他的合作者; 本书包括了他们与 Mordukhovich 的许多合作成果.

本卷翻译的具体分工为: 赵亚莉翻译第 5 章及部分附录, 王东翻译第 6 章, 王炳武翻译前言和第 7 章, 李春负责第 8 章、部分附录的翻译及综合与协调工作. 全书的译稿最后由王炳武做了统一的校对和修改.

翻译过程中对原著中的笔误尽可能做了修正, 其中许多修改来自 Mordukhovich 提供的勘误表; 简单起见译文中没有一一标注. 另外本书中 Mordukhovich 创始的基本非光滑结构现在一般以 Mordukhovich 的名字命名, 比如基本法锥称为 Mordukhovich 法锥, 基本上导数称为 Mordukhovich 上导数, 基本次微分称为 Mordukhovich 次微分等. 但本书译文为忠实于原著, 保持了原来的术语.

译者感谢渤海大学对本书翻译工作的大力支持, 特别感谢张守波副校长和数理学院王志福院长的协助; 感谢大连理工大学张立卫教授对第 5 章译稿给予的建议; 感谢渤海大学单畅教授对第 8 章翻译的帮助. 本书的翻译也得到了 Mordukhovich 教授的支持和鼓励, 在此一并表示感谢.

希望本书下卷中译本的出版能够在上卷的基础上进一步帮助读者了解国际变分分析领域, 特别是其在最优化、控制和经济学上的应用, 对我国的变分分析及其应用的研究和发展有所裨益. 由于时间仓促及译者水平有限, 错误或不当之处在所难免, 恳请读者批评指正.

译 者
2013 年 6 月 9 日

前　言

这就是说, 因为整个宇宙的形式是如此之完美, 事实上由最睿智之造物主所创, 世上没有任何事情之生发不为极大极小原理的光芒所指引.

Leonhard Euler (1744)

Euler[411] (“...nihil omnino in mundo contingint, in quo non maximi minimive ratio quapiam eluceat”) 的这个鲜明观点可作为“变分分析”的最基本原理. 在解决数学和应用科学中的一些可以本不具有变分特性的问题中, 优化和变分方法有各种引人注目的应用, 该原理给出了理论的依据. 众所周知, 优化曾是微分积分发展的主要动力. 事实上, Fermat 通过函数图像的切线斜率而引入的导数概念就是为了解决一个优化问题, 这导出了现在所谓的“Fermat 驻点原理”. Fermat 的这个原理除在优化中的应用以外, 还在微积分最重要的一些结果的证明中举足轻重, 这包括中值定理、隐函数和逆函数定理等. 这样的发展脉络在无限维的情形也可以看到, 其中最速降线问题不仅是变分法的第一个问题, 也同样是整个泛函分析的起点. 特别地, 它催生了无限维空间中的微分及相关领域的各种新概念.

现代变分分析可看为变分法和数学规划的拓展, 它致力于在各种约束下的函数优化以及优化相关问题对于扰动的灵敏性和稳定性. 像离给定点或曲线的位移这样的经典概念已不再紧要, 而问题的逼近或扰动成了关键.

现代变分分析最具代表性的特征之一是其内蕴的非光滑性, 也就是说, 必须处理不可微函数, 具有非光滑边界的集合和集值映射. 而这种非光滑性是自然而然产生的, 它并不仅仅源于优化相关问题 (特别是具有不等式和几何约束的那些) 的初始数据, 更多的来自应用于问题的变分原理及其他优化、逼近和扰动等技术, 而这些问题的数据却可以是光滑的. 事实上, 变分分析框架中经常出现的许多基本概念 (比如距离函数、优化控制中的值函数、极大极小函数、扰动约束和变分系统的解映射等) 都不可避免地是非光滑或是集值的, 这就需要发展涉及“广义微分”的新型分析.

要重点强调的是, 最优控制问题即便是最简单或是最早期的, 也与经典的变分法不同, 它们本质上是非光滑的. 这主要是因为控制函数所具有的逐点约束, 它经常只取离散的值, 比如自动控制中的一些典型问题, 而这些问题是最优控制的一个基本动力. 对变分分析和广义微分的高等方法而言, 最优控制一直是主要的动力源泉和卓有成效的应用领域.

有限维空间中变分分析的要义在 Rockafellar 与 Wets 的书 *Variational Analysis* 中已经得到论述, 但无限维变分分析的应用及发展中所需要的某些概念和工具在有限维理论中是找不到的. 本书的基本目标就是阐明变分分析在有限维和无限维的统一框架下的基本概念和原理, 发展一套与有限维情形同样完美的广义微分的详尽理论, 并提供变分理论在很多领域中广泛而有意义的应用, 这些领域包括约束优化与均衡、灵敏性与稳定性分析、常微分方程、泛函微分方程、偏微分方程的控制理论, 某些选题涉及了力学与经济模型.

变分分析及其应用的核心是广义微分理论. 本书利用对偶空间几何方法, 系统地建立了一套广义微分理论. 它是围绕着"极点原理"展开的, 该原理可以看做经典的凸集分离定理在非凸情形的局部变分对应版本, 它能够处理非凸的集合、集值映射和增广实值函数的类导数结构 (分别是法锥、上导数和次微分). 这些结构是直接在对偶空间中定义的. 因其取值是非凸的, 它们不能由原空间中的类导数结构 (比如切锥和方向导数) 生成. 但是, 基本非凸结构却享有详尽的分析法则, 并远远优于其在原空间中的或凸值的类似结构. 与原空间中的结构相比, 在对偶空间中能促成更多的和谐与美. 从某种意义上讲, 对上面引用的 Euler 基本观念中的完美性, 对偶观点事实上的确达到了这一要求.

在此方向可看到, 对偶结构 (乘子、伴随轨线、影子价格等) 一直就是变分理论及其应用的核心, 特别是在变分法、数学规划、最优控制和经济模型等主要最优条件的表述中. 在原空间中使用最优解的变分只能看做是推导必要最优条件的一个捷径, 这是因为, 在凸和光滑的情形, 原空间中或对偶空间中的类导数结构是等价的, 所以"原空间"方法并不会受到限制. 在现代变分分析中就不一样了, 因为即使在原空间中使用非凸的局部逼近 (比如切锥), 通过对偶而得到的法锥和次微分也不可避免地是凸的. 对偶结构的这种凸性在理论和应用上都导致很大的限制. 进一步, 有很多情形, 特别是本书中要指出的那些, 原空间逼近方法在变分分析中根本用不上, 而利用对偶结构却可得到完备的结果. 当然, 切向生成的或原空间中的结构在变分分析的其他方面还是有重要作用的, 特别是在有限维的情形, 可以通过在所研究的点附近取极限而重建这些非凸的对偶结构. 作为例子, 请见前面提到的 Rockafellar 与 Wets 的书 [1165].

在本书所录的文献中, 特别建议读者参考如下专著: Aubin 与 Frankowska[54], Bardi 与 Capuzzo Dolcetta[85], Beer[92], Bonnans 与 Shapiro[133], Clarke[255], Clarke, Ledyaev, Stern 与 Wolenski[265], Facchinei 与 Pang[424], Klatte 与 Kummer[686], Vinter[1289]. 这个领域发展很快, 对本书中未考虑的变分分析的一些重要方面及应用, 请参阅每章最后给出的评注. 特别强调同时出版的具有互补性的 Borwein 与 Zhu 的专著 [164], *Techniques of Variational Analysis*, 它介绍了本书中没有的一些现代变分分析的基本技巧, 囊括了一些重要的理论和应用.

放在读者面前的这本书是自成一体的，主要集中了尚未见载于专著中的结果，共两卷八章，然后每章分成节和小节．每章都给出了详尽的评注 (这在本书中扮演着一个特别的角色，其中讨论了基本思想、历史、源动力、各种关系、名词选取、未解决问题等)．给出并讨论了很多关于变分分析各个方面 (本书论及或未论及的) 的文献，包括最初的贡献和近期的发展．尽管没有正式的练习题，大量的注释和例子提供了进一步思考和发展的题材．主要结果的证明是完整的，但也预留了一些空间，以补足细节，研究特例及导出一些推广，这时书中经常会给出一些提示．

第一卷“基本理论”包括 4 章，主要涵盖广义微分的基本结构、基本的极点原理和变分原理、完备的广义微分法则以及非线性分析基本性质的完整对偶刻画及其在约束与变分系统灵敏性分析上的应用，这些性质相关于 Lipschitz 稳定性和度量正则性．

第 1 章讨论一般 Banach 空间中的广义微分理论．基本法锥、次微分和上导数是在对偶空间中直接定义的，这涉及更原始的 Fréchet 类型的 ε- 法向量和 ε- 次微分并取弱* 极限而得到．该章指出了这些结构在 Banach 空间中的各种优良特性，此处使用 ε- 法锥是很关键的．这些性质 (包括一阶与二阶微分法则、有效表示、变分描述、距离函数的次导数微分法、Lipschitz 稳定性和度量正则性的上导数必要条件等) 大多收在本章．这里也定义了并开始研究所谓的“序列法紧性”(SNC)，它是集合、集值映射和增广实值函数的性质，在有限维空间中是自动成立的，但却是无限维空间变分分析及其应用的要素之一．

第 2 章细致研究了变分分析中的“极点原理”，它是本书的主要工具．这里首先利用“度量逼近”方法通过一个光滑罚函数过程给出了有限维空间中极点原理的直接变分证明．接着用无限维空间中的变分技巧在具有光滑 Fréchet 范数的 Banach 空间中证明了它，然后利用可分约化，将其推广到 Asplund 空间．Asplund 空间在 Banach 空间几何理论中有很细致的研究，它包括所有的自反空间，以及具有可分对偶的空间．这种空间对本书中发展的变分分析理论和应用起着显著的作用．在这章中，还建立了 (几何) 极点原理和 (解析) 变分原理的关系，这包括传统的形式和改进的形式．应用所得结果，给出了 Asplund 空间的一些新的变分刻画和基本广义微分结构在 Asplund 空间中一些类似于有限维空间中的有用表示．最后，这一章还讨论了恰当 Banach 空间上极点原理的抽象版本，它由以公理定义的法锥和次微分结构给出．对一些特殊的结构，还简要给出了更多的细节．

第 3 章是本书中建立的广义微分理论的基石，它涵盖了 Asplund 空间中基本法锥、次导数和上导数的完备分析法则．该章把主要精力放在了在所考虑点极限结构的点基法则，这既体现在假设中，也体现在结论中．这是因为，点基结果在应用中是至关重要的．本章中给出的有些结果在有限维中似乎也是新的，而整体上这些结果在 Asplund 空间中达到了与有限维空间一样的完美和广泛．区分有限维和无限

维的要点在于在无限维空间中需要足够的紧性, 这在有限维空间中是不需要的. 这里所需的紧性由前面提到的 SNC 性质提供, 这包括在这些微分法则的假设中, 同时也提出了在集合和映射各种运算下 SNC 性质本身分析法则的需求. 这种 SNC 分析法则的缺失是广义微分理论在无限维空间中成功应用的主要羁绊, 这些无限维应用问题包括本书中的优化、稳定性和最优控制. 本章中包括了对这些应用具有决定性意义的广泛的 SNC 法则.

第 4 章详尽研究了集值映射的 Lipschitz 性质、度量正则性和线性开性/覆盖性, 及其在参数约束和变分系统灵敏性上的应用. 首先基于前面建立的变分原理和广义微分理论, 证明了第 1 章给出的这些基本性质在一般 Banach 空间中的那些必要上导数条件, 在 Asplund 空间中就成了这些性质的完整刻画. 进一步, 由变分方法得到计算相应模的可验证的确切公式. 接下来, 利用广义微分法则和 SNC 法则, 提供了这些结果在参数约束和变分系统灵敏性与稳定性上的应用, 这些系统由可行解和最优解的扰动集合给出, 而这些解源于优化与均衡、隐函数、互补条件、变分和半变分不等式等问题以及一些力学系统.

第二卷“应用”也包括 4 章, 主要研讨变分分析基本原理与发展的广义微分理论在许多方面的应用, 这包括约束优化与均衡、常微和分布参数系统的最优控制、福利经济模型等.

第 5 章涉及约束优化和均衡问题, 这里初始数据可能是非光滑的. 即使初始数据是光滑的, 基于极点原理/变分原理与广义微分理论的变分分析的高等方法对约束问题也是非常有用的, 这是因为在应用罚函数、逼近和扰动技术时, 非光滑性会自然而然地产生. 这里的基本目标是, 在有限维和无限维空间中, 对各种类型约束问题导出必要最优和次最优条件. 值得注意的是, 后面的这种次最优条件并不需要假设最优解的存在性 (这在无限维空间中有特别重要的意义), 但保证了“几乎”最优的解“几乎”满足必要最优条件. 这种条件的意义在最优化中似乎被低估了. 除了考虑通常类型的约束问题, 该章还认真研究了一类相当新的问题, 即均衡约束数学问题 (MPEC) 与均衡约束均衡问题 (EPEC), 这些问题具有内蕴的非光滑性, 由广义微分理论可以给出完整的分析. 最后, 该章表述了某些线性次极点和线性次最优的概念, 使得上面以通常概念导出的必要最优条件在新的情形变成了充分必要的.

第 6 章开始研究“动态最优化”和“最优控制”. 正如前面提到的, 这是建立新式变分分析的主要动力之一. 这章主要处理由常微分动力系统控制的最优控制问题, 其状态空间可能是无限维的. 本章的第一部分主要致力于由约束“微分包含”控制的发展系统的 Bolza 类型问题. 这样的模型包括了更常见的由参数发展方程控制的控制系统, 其中的控制区域一般来说依赖于状态变量. 后者不允许使用控制变分来导出必要最优条件. 该章建立了“离散逼近方法”, 它显然在数值分析上有意义, 但在本书中主要是用作一个直接的工具来导出连续时间系统的最优条件, 这

是通过把离散时间量取极限而得到的. 用这个办法, 很强地基于广义微分理论和 SNC 分析法则, 得到了无穷维空间上非凸微分包含的必要最优条件, 这些条件有一般 Euler-Lagrange 的形式, 并由基本微分结构给出.

第 6 章第二部分处理一般 Banach 空间中的约束最优控制系统, 它由光滑动力学的常微分发展方程所控. 与上述微分包含相比, 这类问题具有本质上不同的性质. 本章中两个部分得到的结果 (和所用的方法) 一般来说是独立的. 这里的另一条主线涉及到非凸控制系统最大值原理在离散逼近下的稳定性. 这里建立的"近似最大值原理"是一个有点令人惊讶的结果, 它对连续时间和离散时间控制系统的关系, 包括定量的与定性的, 都有积极的意义.

第 7 章继续研究变分分析高等方法在最优控制问题上的应用, 这里考虑的系统是分布参数的. 该章首先考察了一类一般的"遗传系统", 其动态约束由时滞微分包含和线性代数方程描述. 一方面, 这类控制系统很有意思且尚无太多的研究, 它可以看做"中立型泛函微分包含"变分问题的一个特例, 其中包含系统的时滞不仅存在于状态变量, 也存在于速度变量. 另一方面, 这类系统相关于这样的微分代数系统, 其"慢"和"快"变量之间有线性联系. 利用离散逼近方法和广义微分理论的基本工具, 这里建立了离散逼近的一个强变分收敛性/稳定性, 并导出了连续时间系统广泛的最优条件, 其中包括 Euler-Lagrange 形式和 Hamilton 形式.

第 7 章余下的部分研究了"偏微分方程"控制的最优控制问题, 它具有逐点的控制和状态变量. 这里把主要精力放在由"抛物"和"双曲"方程描述的发展系统, 其控制函数作用于 Dirichlet 和 Neumann 边界条件. 这样的"边界控制"问题在 PDE 最优控制中是最有挑战性和最少被研究的, 特别是具有逐点状态约束的情形. 该章利用现代变分分析的近似和扰动方法, 证明了变分收敛性并导出了这样的 PDE 系统各种控制问题的必要最优条件, 其中包括特定摄动下的极小极大控制.

本书最后的第 8 章是关于变分分析在"经济模型"中的应用的. 这里的主要课题是"福利经济学", 所考虑的是一般的非凸情形, 并具有无限维的商品空间. 这类重要的竞争均衡问题得到了经济学家和数学家的重视, 特别是最近, 实际应用中非凸性变得越来越关键. 本书中发展的变分分析方法, 特别是极点原理, 为这种模型中的 Pareto 最优分配和相关联的价格均衡提供了足够的研究工具. 这里由变分分析和广义微分这些工具推导出所谓的"福利经济学第二基本定理"在非凸情形很广泛的扩展, 它以非凸集合广义法向量的极小组合描述了边际均衡价格. 特别地, 该章的方法和广义法向量的变分刻画给出了市场均衡的新经济解释, 这用到了"非线性边际价格", 其在非凸模型中的角色类似于凸模型中通常 Arrow-Debreu 类型的线性价格.

本书包含一个记号表, 这对两卷是一样的, 对每一卷还有一个详尽的名词索引. 由此索引, 读者不但可以找到该概念/记号在书中首次出现的页码, 也能找出其在

书中各处的进一步讨论及应用.

另外, 书中所有的陈述 (包括定义、定理、引理、命题、推论、例子和注解) 都加了题目. 这些陈述在每一章中按顺序加了序号, 比如在第 5 章中, 例 5.3.3 在定理 5.3.4 前面, 该定理后面则是推论 5.3.5. 为方便读者, 所有这些陈述以及加了序号的注释都罗列在每卷后面的陈述表里. 值得一提的是, 缩略词 (依字母序) 也列在名词索引里. 对所用的记号, 书中一般的原则是以小写希腊字母表示数和 (增广) 实值函数, 以小写拉丁字母表示向量和单值映射, 以大写希腊和拉丁字母表示集合和集值映射.

本书的记号和术语大体上和 Rockafellar 与 Wets 的书 [1165] 中的是一样的. 在全书中都尽量区分概念是定义“在”一个点和一个点“附近”, 后者显示了对扰动的鲁棒性/稳定性, 这在本书中的大部分主要结果是很关键的.

本书附录了丰富的文献 (如果可能都是英文的), 这在上下卷里是一样的. 所录文献反映了各种研究课题和许多研究人员的贡献. 这些文献或多或少主要是在每章的评注里讨论的. 读者可以根据作者的评注在文献中找到更多的信息.

本书主要面向数学科学的研究人员和研究生, 这首先包括那些对如下领域感兴趣的: 非线性分析、最优化、均衡理论、控制论、泛函分析、常/偏微分方程、泛函微分方程、连续介质力学和数理经济. 作者也期望本书会对涉及变分分析的研究及应用的更广泛的研究人员、实际应用者、研究生有用, 相关领域可能包括运筹、统计、力学、工程、经济和其他应用科学.

书中的某些部分已被作者用于州立韦恩大学的变分分析、最优化和最优控制等研究生课程. 书里的基本材料也曾被集成到最近几年作者在许多学校和学术会议所作的讲座里.

致　谢

首先感谢 Terry Rockafellar, 他多年来一直鼓励我写这样一本书, 并在写作过程的每个时期都给予了建议和支持.

特别感谢 Rafail Gabasov, 我的博士论文导师, 我从他那里学到了最优控制和更多其他知识; 感谢 Alec Ioffe, Boris Polyak 和 Vladimir Tikhomirov, 他们发现了我起初在非光滑分析和优化上的工作并给予了有力的支持; 感谢 Sasha Kruger, 我的第一个研究生和合作者, 我们从那时开始了令人振奋的广义微分之旅; 感谢 Jon Borwein 和 Marián Fabian, 从他们那里学到了泛函分析的高深理论和 Asplund 空间的美妙; 感谢 Ali Khan, 他令人鼓舞的工作和热情促进了笔者在经济模型理论上的研究; 感谢 Jiři Outrata, 他的推动和影响使我在均衡和力学问题的兴趣与日俱增, 他并满怀热情地把本书中的广义微分结构具体用于最优化及应用的各个领域; 感谢 Jean-Pierre Raymond, 他使我在现代偏微分方程理论上获益匪浅.

在成书过程中, 我很高兴与许多同事和朋友讨论了本书的许多方面和结果. 除了上面提到的以外, 特别感谢 Zvi Artstein, Jim Burke, Tzanko Donchev, Asen Dontchev, Joydeep Dutta, Andrew Eberhard, Ivar Ekeland, Hector Fattorini, René-Henrion, Jean-Baptiste Hiriart-Urruty, Alejandro Jofré, Abderrahim Jourani, Michal Kočvara, Irena Lasiecha, Claude Lemaréchal, Adam Levy, Adrian Lewis, Kazik Malanowski, Michael Overton, Jong-Shi Pang, Teemu Pennanen, Steve Robinson, Alex Rubinov, Andrzej Świech, Michel Théra, Lionel Thibault, Jay Treiman, Hector Sussmann, Roberto Triggiani, Richard Vinter, Nguyen Dong Yen, George Yin, Jack Warga, Roger Wets 和 Jim Zhu, 感谢他们在成书的这些年中的有价值的建议和卓有成效的交流.

衷心感谢国家科学基金 (NSF) 对笔者研究工作的持续支持.

前面提到, 书中的材料曾被用于讲授变分分析和最优化等高等课程, 参加人员主要包括我的博士生和合作者. 我很感谢他们的贡献, 这特别使我能改进讲义和这本专著. 提供特别帮助的包括: Glenn Malcolm, Nguyen Mau Nam, Yongheng Shao, Ilya Shvartsman 和 Bingwu Wang. 有用的反馈和文字修正也来自 Truong Bao, Wondi Geremew, Pankaj Gupta, Aychi Habte, Kahina Sid Idris, Dong Wang, Lianwen Wang 和 Kaixia Zhang.

我非常感谢 Springer 出版社友好的工作人员, 感谢他们在本书准备和出版过程中的大力支持. 特别的感谢给予数学执行编辑 Catriona Byrne, 应用数学高级编

辑 Achi Dosajh, 助理数学编辑 Stefanie Zoeller 和来自计算机科学编辑部的 Frank Holzwarth.

我感谢我的小女儿 Irina, 感谢她对我的书的兴趣并不厌其烦地回答我在英语上的问题. 我也感谢我的小狮子狗 Wuffy, 感谢它和我分享了长期工作的这些日子. 在所有这些谢意之上, 我无法用足够的话感谢我的妻子 Margaret, 感谢她从我们在 Minsk 高中的日子到现在和我分享的一切.

Boris S. Mordukhovich
Ann Arbor, Michigan
2005 年 8 月

目　录

第 5 章 约束最优化与均衡

这一章研究前面章节建立的变分分析和广义微分的基本工具在可能具有非光滑数据的约束最优化与均衡问题中的应用. 实际上这是个双向过程, 因为由前述, 最优化思想显然也处于变分分析的核心, 特别值得一提的是在有限和无限维空间中所考虑的法向量和次梯度的变分描述; 更多细节参见定理 1.6, 1.1.4 小节和定理 1.88. 而且, 本书分析的主要工具 —— 极点原理 —— 本身给出了集合极性的必要条件, 这些必要条件是在第 2~4 章建立的关于广义微分分析法则和 Lipschitz 稳定性及度量正则性的相关刻画等基本结果的核心.

本章的首要目标是建立无限维空间中约束最优化与均衡等各种问题的必要最优性和次优性条件. 注意到后一种类型的结果 (次优性) 在没有强加精确最优解存在的假设下保证“几乎”最优解“几乎”满足最优性的必要条件, 这在无限维空间中是非常重要的. 本章由泛函和几何约束下的数学规划问题开始, 接着研究多目标最优化、极小极大问题与均衡约束、一些推广极性的概念等各种问题. 本章分析的关键工具是极点原理与它的变体, 以及广义微分分析法则, 其中 SNC 分析法则扮演了一个关键的角色, 这在无限维空间中约束最优化与均衡问题的应用中是至关重要的.

5.1 数学规划的必要条件

本节涉及具有算子、泛函和几何约束的数学规划的一般问题的一阶必要最优性和次优性条件. 这些条件的不同形式随加在原始数据上的假设类型而定. 首先考察 Banach 和 Asplund 空间上只具有由任意非空子集给出的几何约束的最优化问题.

5.1.1 具有几何约束的极小化问题

给定 Banach 空间 X 中的在参考点有限的函数 $\varphi\colon X \to \overline{\mathbb{R}}$ 和一个非空子集 Ω, 考虑下面的具有几何约束的极小化问题:

$$\min \quad \varphi(x) \quad \text{s.t.} \quad x \in \Omega \subset X. \tag{5.1}$$

约束最优化问题 (5.1) 显然等价于无约束问题:

$$\min \quad \varphi(x) + \delta(x; \Omega), \quad x \in X,$$

这里指示函数 $\delta(\cdot;\Omega)$ 在约束外的点上加了一个“无限惩罚”. 这样, 给定 (5.1) 的一个局部最优解 $\bar{x}$, 根据命题 1.114 的广义 Fermat 法则可得

$$0 \in \widehat{\partial}\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}) \subset \partial\big(\varphi + \delta(\cdot;\Omega)\big)(\bar{x}). \tag{5.2}$$

根据初始数据 (φ, Ω), 为使 (5.2) 转变成有效必要最优性条件, 需利用针对 $\varphi+\delta(\cdot;\Omega)$ 的次微分和法则. 如果 φ 在 $\bar{x}$ 是 Fréchet 可微的, 那么这个方向上最简单的结果可由命题 1.107(i) 中的和法则如下得到.

命题 5.1(具有 Fréchet 可微费用函数的约束问题的必要条件)　设 $\bar{x}$ 是 Banach 空间 X 中问题 (5.1) 的局部最优解. 假设 φ 在 $\bar{x}$ 是 Fréchet 可微的. 则

$$-\nabla\varphi(\bar{x}) \in \widehat{N}(\bar{x};\Omega), \quad -\nabla\varphi(\bar{x}) \in N(\bar{x};\Omega).$$

证明　对 (5.2) 中的第一个包含关系应用命题 1.107(i), 并利用关系 $\widehat{\partial}\delta(\bar{x};\Omega) = \widehat{N}(\bar{x};\Omega)$, 可得 $-\nabla\varphi(\bar{x}) \in \widehat{N}(\bar{x};\Omega)$. 由包含关系 $\widehat{N}(\bar{x};\Omega) \subset N(\bar{x};\Omega)$, 亦知第二个必要条件成立. △

如果 φ 在 $\bar{x}$ 不是 Fréchet 可微的, 对类 Fréchet 次梯度结构则不能用上面的方法, 因为类 Fréchet 次梯度结构即使在有限维空间中也不具有令人满意的分析法则. 而对本书中的基本结构 $\partial\varphi$ 和 $N(\cdot;\Omega)$ 而言, 情形则完全不同, 因为它们在 Asplund 空间的一般非光滑情形下具有完整的分析法则. 在开展这个方向的研究之前, 先给出一个关于极小化问题 (5.1) 的出人意料的上次微分必要条件, 这对某些特殊的函数 φ 是很有效的. 这些必要最优性条件推广了命题 5.1 中的相应条件, 实际上在证明中简化为这些条件, 这是由于定理 1.88(i) 中 Fréchet 次梯度的变分描述可以应用于 (1.52) 中定义的 Fréchet 上次微分 $\widehat{\partial}^+\varphi(\bar{x})$.

命题 5.2(几何约束下局部极小值的上次微分条件)　设 $\bar{x}$ 是 Banach 空间 X 中极小化问题 (5.1) 的局部最优解, 这里 $\varphi\colon X \to \overline{\mathbb{R}}$ 在 $\bar{x}$ 有限. 则有包含关系

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset \widehat{N}(\bar{x};\Omega), \quad -\widehat{\partial}^+\varphi(\bar{x}) \subset N(\bar{x};\Omega). \tag{5.3}$$

证明　只需证 (5.3) 中的第一个包含关系成立即可. 当 $\widehat{\partial}^+\varphi(\bar{x}) = \varnothing$ 时这个包含关系是平凡的. 假设 $\widehat{\partial}^+\varphi(\bar{x}) \neq \varnothing$, 取 $x^* \in \widehat{\partial}^+\varphi(\bar{x}) = -\widehat{\partial}(-\varphi)(\bar{x})$. 把定理 1.88(i) 应用于 $\widehat{\partial}(-\varphi)(\bar{x})$ 中的 Fréchet 次梯度 $-x^*$, 可找到一个函数 $s\colon X \to \mathbb{R}$ 满足: $s(\bar{x}) = \varphi(\bar{x})$ 且对任意的 $x \in X$ 有 $s(x) \geqslant \varphi(x)$, 使得 $s(\cdot)$ 在 $\bar{x}$ 是 Fréchet 可微的, 并且 $\nabla s(\bar{x}) = x^*$. 对 $\bar{x}$ 附近的所有 $x \in \Omega$, 有

$$s(\bar{x}) = \varphi(\bar{x}) \leqslant \varphi(x) \leqslant s(x).$$

因此 $\bar{x}$ 是具有 Fréchet 可微目标函数的约束最优化问题

$$\min \quad s(x) \quad \text{s.t.} \quad x \in \Omega$$

的局部最优解. 把命题 5.1 应用于上面的最优化问题, 得 $-x^* \in \widehat{N}(\bar{x};\Omega)$, 从而得 (5.3) 并完成了命题的证明. △

当 φ 在 $\bar{x}$ 是 Fréchet 可微的, 此时由于 $\widehat{\partial}^+\varphi(\bar{x}) = \{\nabla\varphi(\bar{x})\}$, 命题 5.2 的结果简化为命题 5.1 中的包含关系 $-\nabla\varphi(\bar{x}) \in \widehat{N}(\bar{x};\Omega)$. 当 φ 是凹的, 且在 $\bar{x}$ 附近连续, 则满足命题 5.2 假设的一类有趣的最优化问题包含凹极小化问题. 此时 $\widehat{\partial}^+\varphi(\bar{x})$ 与凸分析中的 (非空) 上次微分相同. 如果 X 是 Asplund 空间, 当 φ 在 $\bar{x}$ 附近是 Lipschitz 连续的, 且在 $\bar{x}$ 是上正则的, 即 $\widehat{\partial}^+\varphi(\bar{x}) = \partial^+\varphi(\bar{x})$ 时, 有 $\widehat{\partial}^+\varphi(\bar{x}) \neq \varnothing$. 实际上, 根据推论 2.25, 有 $\partial^+\varphi(\bar{x}) = -\partial(-\varphi)(\bar{x}) \neq \varnothing$. 注意到最后一类函数, 除了包含严格可微函数和凹连续函数外, 还包含对许多应用来说非常重要的所谓半凹函数; 参见 5.5.4 小节中更多的讨论.

值得注意的是, 若有条件 $\widehat{\partial}^+\varphi(\bar{x}) = \varnothing$, 则 (5.3) 中的包含关系平凡成立. 若不考虑约束且 φ 在 $\bar{x}$ 不是 Fréchet 可微的, 则该条件对 (5.1) 来说是一个易于验证的必要最优条件. 事实上, 由于在局部极小值点 $0 \in \widehat{\partial}\varphi(\bar{x}) \neq \varnothing$, 故根据命题 1.87, $\widehat{\partial}^+\varphi(\bar{x})$ 必是空集. 然而, 它是一个平凡的必要条件, 对约束最小化问题它不带有多少信息. 下面用费用函数 φ 的基本和奇异下次梯度表示的下次微分条件对约束极小化来说是更常见的.

命题 5.3(几何约束下局部极小值的下次微分条件) 设 $\bar{x}$ 是极小化问题 (5.1) 的局部最优解, 这里 Ω 是局部闭的, φ 在 $\bar{x}$ 附近是 l.s.c. 的, 且 X 是 Asplund 空间. 假设

$$\partial^\infty\varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) = \{0\}, \tag{5.4}$$

并且假设 Ω 在 $\bar{x}$ 是 SNC 的, 或者 φ 在 $\bar{x}$ 是 SNEC 的 (如果 φ 在 $\bar{x}$ 附近是局部 Lipschitz 的, 那么这些假设都成立). 则有

$$\partial\varphi(\bar{x}) \cap \big(-N(\bar{x};\Omega)\big) \neq \varnothing, \quad \text{i.e.} \ \ 0 \in \partial\varphi(\bar{x}) + N(\bar{x};\Omega). \tag{5.5}$$

证明 把定理 3.36 中次微分和法则应用于 (5.2) 中和的基本次微分即得. △

注 5.4(关于局部极小值的上次微分和下次微分条件比较) 值得注意的是, 尽管命题 5.3 有更广泛的应用性, 但对特殊类非光滑问题, 命题 5.2 中的上次微分条件可以给出更强的必要性结果, 即使对有限维空间中 Lipschitz 函数 φ 的情形亦如此. 特别地, 根据定理 1.93, 对凹连续函数 φ, 有

$$\partial\varphi(\bar{x}) \subset \partial^+\varphi(\bar{x}) = \widehat{\partial}^+\varphi(\bar{x}) \neq \varnothing.$$

因此把 (5.3) 中的第二个包含关系 (它甚至比其中的第一个包含关系更弱) 与 (5.5) 中的包含关系相比, 可看到命题 5.2 中的必要条件要求集合 $\widehat{\partial}^+\varphi(\bar{x})$ 中的“任意”元素 x^* 必属于 $-N(\bar{x};\Omega)$, 而不是命题 5.3 中更小集合 $\partial\varphi(\bar{x})$ 中的“某个”元素 x^*

属于 $-N(\bar{x};\Omega)$. 这表明局部极小值的上次微分必要条件在应用时可能大大优于上面的下次微分条件. 下面通过一个简单的例子来说明:

$$\min \varphi(x) := -|x| \quad \text{s.t.} \quad x \in \Omega := [-1, 0] \subset \mathbb{R}.$$

显然 $\bar{x} = 0$ 不是此问题的一个最优解. 然而, 根据下次微分必要条件 (5.5) 它不能被排除, 因为

$$\partial\varphi(0) = \{-1, 1\}, \quad N(0;\Omega) = [0, \infty), \quad -1 \in -N(0;\Omega).$$

另一方面, 此时 (5.3) 中的两个上次微分必要条件是一样的, 但对 $\bar{x} = 0$ 都不成立, 这是因为

$$\widehat{\partial}^{+}\varphi(0) = [-1, 1], \quad [-1, 1] \not\subset N(0;\Omega).$$

因此该例问题中 $\bar{x} = 0$ 的非最优性可由命题 5.2 确认, 而不是命题 5.3.

注意到关于两个凸函数之差 (即所谓的 DC- 函数, 这在很多应用中是重要的) 的一类极小化问题能等价地简化为求凸约束下的凹函数的极小化问题; 更多研究和讨论, 可见 Horst, Pardalos 和 Thoai [583].

还注意到, 当 φ 在 $\bar{x}$ 是上正则的, 且在该点附近是 Lipschitz 连续的, 而 X 是 Asplund 的, 则有 φ 在 $\bar{x}$ 的 Clarke 广义梯度与 Fréchet 上次微分之间的关系

$$\partial_C\varphi(\bar{x}) = \text{cl}^*\widehat{\partial}^{+}\varphi(\bar{x}).$$

事实上, 由广义梯度的对称性 (2.71) 及其定理 3.57(ii) 中的基本次微分的表示可得. 而且, 如果 X 是 WCG 的, 上面的弱 * 闭包算子是多余的; 见定理 3.59(i). 因此, 在这种情形下用 Clarke 次微分取代命题 5.3 中的基本次微分 $\partial\varphi(\bar{x})$, 所得的下次微分结果远远弱于命题 5.2 中的上次微分条件 (其中 $\widehat{\partial}^{+}\varphi(\bar{x}) = \partial_C\varphi(\bar{x})$).

在变分理论和应用 (特别地, 在最优控制中的应用) 的许多领域中, 几何约束通常由集合的交集给出; 例如, 参见 5.2 节和第 6 章. 基于上面关于问题 (5.1) 的结果和关于交集的基本法向量的分析法则, 可得到具有多个几何约束的最优化问题的必要最优性条件. 为此, 在上次微分条件的情形下, 可利用 (5.3) 中的第二个包含关系. 第一个包含关系对交集得不出有价值的基于点的结果, 这是由于没有足够的 Fréchet 法向量的分析法则.

下面用下和上次微分两种形式给出一般的结果. 为简单起见, 只考虑在 Asplund 空间的乘积空间中给出的几何约束是两个集合的交集的情形. 在定理 5.5 中, 将应用 3.1.1 小节中引入的规范和 PSNC 条件; 对这些条件的讨论见该节.

定理 5.5(几何约束为交集的局部极小值) 设 $\bar{x}$ 是 $\Omega = \Omega_1 \cap \Omega_2$ 时问题 (5.1) 的局部最优解, 这里集合 $\Omega_1, \Omega_2 \subset \prod_{j=1}^{m} X_j$ 在 $\bar{x}$ 附近是局部闭的, 且空间 X_j 是 Asplund 空间. 则下面的结论成立:

(i) 假设系统 $\{\Omega_1, \Omega_2\}$ 在 $\bar{x}$ 满足极限规范条件. 给定 $J_1, J_2 \subset \{1, \cdots, m\}$ 且 $J_1 \cup J_2 = \{1, \cdots, m\}$, 还假设 Ω_1 在 $\bar{x}$ 相对于 J_1 是 PSNC 的, Ω_2 在 $\bar{x}$ 相对于 J_2 是强 PSNC 的. 则有

$$-\widehat{\partial}^+\varphi(\bar{x}) \subset N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2).$$

(ii) 除了 (i) 中的假设外, 假设 φ 在 $\bar{x}$ 附近是 l.s.c. 的, 且在该点 SNEC, 并且设

$$\big(-\partial^\infty\varphi(\bar{x})\big) \cap \big[N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2)\big] = \{0\} \tag{5.6}$$

(如果 φ 在 $\bar{x}$ 附近是 Lipschitz 连续的, 则所有的额外假设都满足). 则有

$$0 \in \partial\varphi(\bar{x}) + N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2). \tag{5.7}$$

(iii) 假设 φ 在 $\bar{x}$ 附近是 l.s.c. 的, Ω_1 和 Ω_2 在该点都是 SNC 的, 并且假设规范条件

$$\Big[x^* \in \partial^\infty\varphi(\bar{x}), x_1^* \in N(\bar{x}; \Omega_1), x_2^* \in N(\bar{x}; \Omega_2), x^* + x_1^* + x_2^* = 0\Big] \Longrightarrow x^* = x_1^* = x_2^* = 0 \tag{5.8}$$

成立, 则有 (5.7).

证明 为证 (i), 应用命题 5.2, 然后对 (5.3) 中的基本法锥 $N(\bar{x}; \Omega)$ 应用定理 3.4 中的交法则, 就有

$$N(\bar{x}; \Omega) = N(\bar{x}; \Omega_1 \cap \Omega_2) \subset N(\bar{x}; \Omega_1) + N(\bar{x}; \Omega_2), \tag{5.9}$$

从而得到 (i) 中的上次微分包含关系.

结论 (ii) 通过把 (5.9) 代入 (5.4) 和 (5.5), 在关于 φ 的 SNEC 假设下利用命题 5.3 及定理 3.4 中的交法则即得. 最后回顾在 $\bar{x}$ 附近局部 Lipschitz 连续的任意函数 φ 根据推论 1.69(参见定义 1.116 后面的讨论) 在 $\bar{x}$ 是 SNEC 的, 及推论 1.81, 则有 $\partial^\infty\varphi(\bar{x}) = \{0\}$.

最后来证 (iii). 证明是应用 Ω 为 SNC 情形的命题 5.3, 这需要由 Ω_1, Ω_2 和 φ 来得出 Ω 是 SNC 的以及命题 5.3 中的其他条件. 如果 Ω_i 在 $\bar{x}$ 都是 SNC 的, 并且满足规范条件

$$N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big) = \{0\}, \tag{5.10}$$

则根据推论 3.81, 交集 $\Omega = \Omega_1 \cap \Omega_2$ 在该点也是 SNC 的. 另外 (5.10) 也保证交公式 (5.9) 成立; 参见推论 3.5. 易验证 (5.8) 蕴涵着规范条件 (5.4) 和 (5.10) 都成立. 事实上, (5.10) 恰好由 $x^* = 0$ 时的 (5.8) 可得. 为得 (5.4), 取 $x^* \in N(\bar{x}; \Omega_1 \cap \Omega_2)$,

$-x^* \in \partial^\infty \varphi(\bar{x})$, 且根据推论 3.5, 找到 $x_i^* \in N(\bar{x}; \Omega_i), i = 1, 2$ 满足 $x_1^* + x_2^* = x^*$. 因此 $-x^* + x_1^* + x_2^* = 0$, 从而由 (5.8) 得 $x^* = 0$, 这就结束了定理的证明. △

下面给出定理 5.5 的一个推论, 它统一和简化了有限多几何约束情形下的假设.

推论 5.6(多个几何约束下的局部极小值) 设 $\bar{x}$ 是问题 (5.1) 的一个局部最优解, 其中 $\Omega = \Omega_1 \cap \cdots \cap \Omega_n$, 且每个 Ω_i 在 Asplund 空间 X 中的点 $\bar{x}$ 附近是局部闭的. 假设除一个外其他所有的 Ω_i 在 $\bar{x}$ 都是 SNC 的, 并且假设

$$\left[x_1^* + \cdots + x_n^* = 0, \quad x_i^* \in N(\bar{x}; \Omega_i)\right] \Longrightarrow x_i^* = 0, \quad i = 1, \cdots, n. \tag{5.11}$$

则上次微分必要条件

$$-\widehat{\partial}^+ \varphi(\bar{x}) \subset N(\bar{x}; \Omega_1) + \cdots + N(\bar{x}; \Omega_n)$$

成立. 而且, 如果 φ 在 $\bar{x}$ 附近是 l.s.c. 的, 且在该点是 SNEC 的, 并且如果 (5.11) 替换成更强的规范条件

$$\left[x^* \in \partial^\infty \varphi(\bar{x}), x_i^* \in N(\bar{x}; \Omega_i),\ i = 1, \cdots, n, \quad x^* + \sum_{i=1}^n x_i^* = 0\right]$$

$$\Longrightarrow x^* = x_1^* = \cdots = x_n^* = 0,$$

则下次微分包含关系

$$0 \in \partial \varphi(\bar{x}) + N(\bar{x}; \Omega_1) + \cdots + N(\bar{x}; \Omega_n)$$

成立. 另外, 如果在上面的假设中 φ 在 $\bar{x}$ 的 SNEC 性质替换成所有 $\Omega_1, \cdots, \Omega_n$ 在该点的 SNC 性质, 则上面的必要最优性条件仍成立.

证明 显然规范条件 (5.11) 及除一个 Ω_i 外其他所有的 Ω_i 的 SNC 性质蕴涵着定理 5.5(i) 的假设对两个集合成立, 进而根据数学归纳法, 对 n 个集合也成立; 从而保证交法则

$$N(\bar{x}; \Omega_1 \cap \cdots \cap \Omega_n) \subset N(\bar{x}; \Omega_1) + \cdots + N(\bar{x}; \Omega_n)$$

成立; 请比照推论 3.37. 这就证明了该推论的上次微分必要条件成立. 下次微分条件在 φ 的 SNEC 假设下, 由定理 5.5 的结论 (ii) 及关于所有 Ω_i 是 SNC 的假设下, 由结论 (iii) 根据数学归纳法可得. △

5.1.2 算子约束下的必要条件

本小节将得到推广的数学规划问题的必要最优性条件, 这种推广的数学规划问题除包含几何约束外, 还包含由在无限维空间中取值的集值和单值映射给出的算子

约束. 本节的方法是将该问题简约为只有几何约束的极小化问题, 其中几何约束为两个集合的交集, 这两个集合中有一个集合是某集合在集值或单值映射下的逆像. 然后利用在第 1、3 章中建立的广义微分分析法则的结果, 其中包括保证 SNC 性质成立和保持的有效法则. 用这种方法, 在某些约束规范下 (这些规范保证所谓最优性条件的标准/规范形式), 得到下和上次微分类型的一般必要最优性条件, 以及没有这样规范的必要条件.

考虑下面的约束最优化问题:

$$\min \quad \varphi_0(x) \quad \text{s.t.} \quad x \in F^{-1}(\Theta) \cap \Omega, \tag{5.12}$$

这里, $\varphi_0\colon X \to \overline{\mathbb{R}}, F\colon X \rightrightarrows Y, \Omega \subset X, \Theta \subset Y$, 并且

$$F^{-1}(\Theta) := \{x \in X \mid F(x) \cap \Theta \neq \varnothing\}$$

是集合 Θ 在集值映射 F 下的逆像. 模型 (5.12) 包括许多特殊类最优化问题, 特别地, 包括具有等式和不等式约束的经典的非线性规划问题; 参见 5.1.3 小节.

注意到 (5.12) 可简化为只有几何约束的约束极小化问题, 若几何约束由两个集合 $\Omega_1 = F^{-1}(\Theta), \Omega_2 = \Omega$ 的交集给出. 因此可利用 5.1.1 小节的结果, 然后利用逆像和交集的法锥分析法则以及保持 SNC 性质的分析法则, 这些结果已在第 1、3 章建立. 这样就得到一般问题 (5.12) 标准形式的必要最优性条件, 即非零乘子对应着价值函数 φ_0. 下面先建立极小化问题 (5.12) 的上次微分必要最优性条件.

定理 5.7(算子约束下局部极小值的上次微分条件) 给定问题 (5.12) 的一个局部最优解 $\bar{x}$, 其中 X 和 Y 是 Banach 空间, 则有如下结论成立:

(i) 假设 $\Omega = X, F = f\colon X \to Y$ 在 $\bar{x}$ 是 Fréchet 可微的, 具有满射导数 $\nabla f(\bar{x})$, 又假设 f 在 $\bar{x}$ 是严格可微的, 或它在该点附近是连续的且 $\dim Y < \infty$. 则

$$-\widehat{\partial}^+\varphi_0(\bar{x}) \subset \nabla f(\bar{x})^* \widehat{N}\big(f(\bar{x}); \Theta\big).$$

(ii) 假设 X 是 Asplund 空间, Ω 在 $\bar{x}$ 附近是局部闭的, $F = f\colon X \to Y$ 在 $\bar{x}$ 是严格可微的, 具有满射导数, 并且规范条件

$$\nabla f(\bar{x})^* N\big(f(\bar{x}); \Theta\big) \cap \big(-N(\bar{x}; \Omega)\big) = \{0\}$$

成立. 如果 Ω 或 Θ 分别在 $\bar{x}$ 和 $f(\bar{x})$ 是 SNC 的, 则有

$$-\widehat{\partial}^+\varphi_0(\bar{x}) \subset \nabla f(\bar{x})^* N\big(f(\bar{x}); \Theta\big) + N(\bar{x}; \Omega).$$

(iii) 假设空间 X 和 Y 都是 Asplund 空间, 集合 Ω, Θ 和 gph F 是闭的, 集值映射 $S(\cdot) := F(\cdot) \cap \Theta$ 在 $\bar{x}$ 附近内半紧. 则有

$$-\widehat{\partial}^+\varphi_0(\bar{x}) \subset \bigcup \Big[D_N^* F(\bar{x}, \bar{y})(y^*) \Big| \bar{y} \in S(\bar{x}), y^* \in N(\bar{y}; \Theta)\Big] + N(\bar{x}; \Omega) \tag{5.13}$$

成立, 如果关于 (F, Θ, Ω) 的下列条件之一满足:

(a) Ω 在 $\bar{x}$ 是 SNC 的, 规范条件

$$\bigcup \Big[D_N^* F(\bar{x}, \bar{y})(y^*) \Big| \bar{y} \in S(\bar{x}), y^* \in N(\bar{y}; \Theta) \Big] \cap \big(-N(\bar{x}; \Omega) \big) = \{0\}, \tag{5.14}$$

$$N(\bar{y}; \Theta) \cap \ker \widetilde{D}_M^* F(\bar{x}, \bar{y}) = \{0\}, \quad \forall\, \bar{y} \in S(\bar{x}) \tag{5.15}$$

满足, 且对任意的 $\bar{y} \in S(\bar{x})$, 逆映射 F^{-1} 在 $(\bar{y}, \bar{x})$ 是 PSNC, 或 Θ 在 $\bar{y}$ 是 SNC 的.

(b) 规范条件 (5.14) 和

$$N(\bar{y}; \Theta) \cap \ker D_N^* F(\bar{x}, \bar{y}) = \{0\}, \quad \forall \bar{y} \in S(\bar{x}) \tag{5.16}$$

满足, 且对任意的 $\bar{y} \in S(\bar{x})$, 或者 F 在 $(\bar{x}, \bar{y})$ 是 PSNC 的, 并且 Θ 在 $\bar{y}$ 是 SNC 的; 或者 F 在 $(\bar{x}, \bar{y})$ 是 SNC 的.

证明 为在 Banach 空间情形证明 (i) 成立, 根据命题 5.2 中第一个上次微分条件, 再利用推论 1.15 中计算 $\widehat{N}(\bar{x}; f^{-1}(\Theta))$ 的等式即得.

当 X 是 Asplund 空间 (而 Y 可以是任意 Banach 空间), f 在 $\bar{x}$ 是严格可微的, 且具有满射导数时, 现在来证 (ii) 成立. 方法是在假设 Ω 或 $f^{-1}(\Theta)$ 分别在 $\bar{x}$ 和 $f(\bar{x})$ 是 SNC, 且

$$N\big(\bar{x}; f^{-1}(\Theta)\big) \cap \big(-N(\bar{x}, \Omega) \big) = \{0\}$$

成立时, 取 $\Omega_1 = f^{-1}(\Theta), \Omega_2 = \Omega$, 然后利用定理 5.5 的结论 (i). 当 Ω 在 $\bar{x}$ 是 SNC 的, 由定理 1.17 中的法锥表达式

$$N\big(\bar{x}; f^{-1}(\Theta)\big) = \nabla f(\bar{x})^* N\big(f(\bar{x}); \Theta\big)$$

即得 (ii). 当 Ω 在 $\bar{x}$ 不是 SNC 的, 则需要利用 $f^{-1}(\Theta)$ 在 $\bar{x}$ 的 SNC 性质, 根据定理 1.22, 它等价于 Θ 在 $f(\bar{x})$ 的 SNC 性质. 因此 (ii) 成立.

为证结论 (iii), 当 $\Omega_1 = F^{-1}(\Theta), \Omega_2 = \Omega$ 时利用定理 5.5(i). 然后利用定理 3.8 中一般集值映射的 $N\big(\bar{x}; F^{-1}(\Theta)\big)$ 的上估计, 其中要求空间 X 和 Y 都是 Asplund 空间. 为利用这个定理, 首先注意到在 (iii) 的假设下集合 $F^{-1}(\Theta)$ 在 $\bar{x}$ 附件是局部闭的; 参见定理 3.8 的证明, 其中注意到 $S(\cdot)$ 被假定为在 $\bar{x}$ 附近是下半紧的. 由定理 3.8, 在 (a) 中所作的关于 F 和 Θ 的假设下, 得

$$N\big(\bar{x}; F^{-1}(\Theta)\big) \subset \bigcup \Big[D_N^* F(\bar{x}, \bar{y})(y^*) \Big| \bar{y} \in S(\bar{x}), y^* \in N(\bar{y}; \Theta) \Big].$$

现在, 如果 Ω 被假定为在 $\bar{x}$ 是 SNC 的, 则在规范条件

$$N\big(\bar{x}; F^{-1}(\Theta)\big) \cap \big(-N(\bar{x}; \Omega) \big) = \{0\}$$

下利用定理 5.5 中的上次微分包含关系

$$-\widehat{\partial}^+\varphi_0(\bar{x}) \subset N\big(\bar{x}; F^{-1}(\Theta)\big) + N(\bar{x}; \Omega)$$

得 (5.13) 成立.

如果 Ω 没有被假定为在 $\bar{x}$ 是 SNC 的, 则需应用 $F^{-1}(\Theta)$ 在 $\bar{x}$ 的 SNC 性质, 在 (b) 中所作的假设下, 这一性质可由定理 3.84 保证. 这就完成了定理的证明. △

值得注意的是, 根据命题 1.68, 如果 F 在 $(\bar{x}, \bar{y})$ 附近是类 Lipschitz 的, 则 (b) 中 F 的 PSNC 性质成立. 还注意到如果 X 额外地假定为 Asplund 空间, 而且 Θ 在 $f(\bar{x})$ 附近是局部闭的, 则定理 5.7 中结论 (ii) 的结果简化为该定理中结论 (iii) 的结果.

下面基于定理 5.5 的结论 (ii) 和 (iii), 并且利用定理 5.7 的证明中所用的分析法则, 得到问题 (5.12) 标准形式的下次微分最优性条件.

定理 5.8(算子约束下局部极小值的下次微分条件) 给定问题 (5.12) 的一个局部最优解 $\bar{x}$, 假设 X 是 Asplund 空间, Ω 在 $\bar{x}$ 附近是局部闭的, φ_0 在该点附近是 l.s.c. 的. 则有下列结论成立:

(i) 设 Y 是 Banach 空间, $F = f\colon X \to Y$ 在 $\bar{x}$ 是严格可微的, 具有满射导数 $\nabla f(\bar{x})$. 如果

$$\Big[x^* \in \partial^\infty\varphi_0(\bar{x}), x_1^* \in \nabla f(\bar{x})^* N\Big(f(\bar{x}); \Theta\Big), x_2^* \in N(\bar{x}; \Omega), x^* + x_1^* + x_2^* = 0\Big] \Rightarrow x^* = x_1^* = x_2^* = 0$$

和下列条件之一成立:

(a) φ_0 在 $\bar{x}$ 是 SNEC 的, Ω 在 $\bar{x}$ 是 SNC 的或者 Θ 在 $f(\bar{x})$ 是 SNC 的;

(b) Ω 和 Θ 分别在 $\bar{x}$ 和 $f(\bar{x})$ 都具有 SNC 性质.

则有

$$0 \in \partial\varphi_0(\bar{x}) + \nabla f(\bar{x})^* N(f(\bar{x}); \Theta) + N(\bar{x}; \Omega).$$

(ii) 设 Y 是 Asplund 空间, 集合 Θ 和 gph F 是闭的, 集值映射 $S(\cdot) = F(\cdot) \cap \Theta$ 在 $\bar{x}$ 附近是内半紧的. 则有

$$0 \in \partial\varphi_0(\bar{x}) + \bigcup \Big[D_N^* F(\bar{x}, \bar{y})(y^*) \,\Big|\, \bar{y} \in S(\bar{x}), y^* \in N(\bar{y}; \Theta)\Big] + N(\bar{x}; \Omega), \tag{5.17}$$

如果关于 $(\varphi_0, F, \Theta, \Omega)$ 的下列假设之一成立:

(c) φ_0 在 $\bar{x}$ 是 SNEC 的, 并且

$$\Big[x^* \in \partial^\infty\varphi_0(\bar{x}), x_1^* \in \bigcup \Big[D_N^* F(\bar{x}, \bar{y})(y^*) \,\Big|\, \bar{y} \in S(\bar{x}), y^* \in N(\bar{y}; \Theta)\Big],$$
$$x_2^* \in N(\bar{x}; \Omega), x^* + x_1^* + x_2^* = 0\Big] \Rightarrow x^* = x_1^* = x_2^* = 0. \tag{5.18}$$

另外, 定理 5.7(iii) 中的 (a) 或 (b) 中的假设成立, 其中 (5.14) 被 (5.18) 取代.

(d) Ω 在 $\bar{x}$ 是 SNC 的, 规范条件 (5.16) 和 (5.18) 满足, 对任意的 $\bar{y} \in S(\bar{x})$, 或者 F 在 $(\bar{x}, \bar{y})$ 是 PSNC 的, 且 Θ 在 $\bar{y}$ 是 SNC 的, 或者 F 在 $(\bar{x}, \bar{y})$ 是 SNC 的.

证明 为证结论 (i), 当 $\Omega_1 = f^{-1}(\Theta), \Omega_2 = \Omega$ 时利用定理 5.5(ii). 在 (a) 情形下想要的结果根据定理 5.7(ii) 的证明中法锥 $N(\bar{x}; f^{-1}(\Theta))$ 的表达式即得. 当 φ_0 在 $\bar{x}$ 不是 SNEC 的, 则需要利用所给条件证明 $f^{-1}(\Theta) \cap \Omega$ 在 $\bar{x}$ 的 SNC 性质成立. 由于在 (b) 中所作的假设下集合 $f^{-1}(\Theta)$ 和 Ω 在该点都是 SNC 的, 又由于 $\nabla f(\bar{x})$ 是满射, 所以该交集的 SNC 性质由推论 3.81 可得.

结论 (ii) 的证明是类似的, 即当 $\Omega_1 = F^{-1}(\Theta), \Omega_2 = \Omega$ 时应用定理 5.5(ii) 和定理 5.7 (iii) 的证明中 $N(\bar{x}; F^{-1}(\Theta))$ 的上估计. 这就给出了情形 (c) 下的次微分包含关系 (5.17). 为在情形 (d) 下证明 (5.17) 成立, 注意到集合 $F^{-1}(\Theta)$ 和 Ω 在所给的假设下在 $\bar{x}$ 都是 SNC 的, 根据推论 3.81, 规范条件 (5.18) 确保 $F^{-1}(\Theta) \cap \Omega$ 的 SNC 性质. 这就完成了定理的证明. △

注意到如果空间 Y 是 Asplund 空间, 集合 Θ 是闭的, 则定理 5.8 中的结论 (i) 的结果由结论 (ii) 中的结论直接可得. 然而, 在 (i) 中一般来说不需这些假设. 还注意到当 X 是有限维的, 规范条件 (5.15) 和 (5.16) 相同, 而一般来说 (5.15) 更弱. (5.15) 的主要好处在于, 如果 F 在该点附近是度量正则的, 则它及 F^{-1} 在 $(\bar{y}, \bar{x})$ 的 PSNC 性质总一起成立. 这样就得到定理 5.7 和 5.8 下面的有效推论, 其中费用函数 φ_0 被假设为局部 Lipschitz 的, 以便简化定理 5.8 中的条件.

推论 5.9(度量正则约束下上和下次微分条件) 设 $\bar{x}$ 是问题 (5.12) 的局部最优解. 除定理 5.7(iii) 的假设, 再设 (5.14) 与 Ω 在 $\bar{x}$ 是 SNC 的, 还假设对任意的 $\bar{y} \in S(\bar{x}), F$ 在 $(\bar{x}, \bar{y})$ 附近是度量正则的, 则上次微分条件 (5.13) 成立. 更进一步, 如果 φ_0 在 $\bar{x}$ 附近是局部 Lipschitz 的, 则下次微分条件 (5.17) 也成立.

证明 根据定理 4.18(c) 中度量正则性的上导数刻画, 由定理 5.7(iii) 的情形 (a) 即得上次微分条件 (5.15). 从定理 5.8(ii) 的 (c) 情形中即可得到下次微分条件 (5.17), 只需注意到若 φ_0 是局部 Lipschitz 的, 则它在 $\bar{x}$ 自动是 SNEC 的, 并且在这个假设下 (5.18) 简化为 (5.14). △

如果 F 在极小值点 $\bar{x}$ 处是单值和严格 Lipschitz 的, 则问题 (5.12) 的上次微分必要最优性条件 (5.13) 和下次微分必要最优性条件 (5.17) 有本质的简化. 这是由于在 f 是 w^*- 严格 Lipschitz 的假设下在定理 3.28 中建立的基本上导数的标量化公式. 注意到对 Asplund 空间之间的映射来说, 根据命题 3.26, 定义 3.25 中的严格 Lipschitz 的概念与 w^*- 严格 Lipschitz 概念是等价的.

推论 5.10(严格 Lipschitz 约束下上和下次微分条件) 设 $\bar{x}$ 是 Asplund 空间 X 和 Y 中问题 (5.12) 的局部解, 这里 $F = f\colon X \to Y$ 是单值的, 且在 $\bar{x}$ 是严格 Lipschitz 的. 在定理 5.7(iii) 与定理 5.8(ii) 的相应假设下, 其中 $S(\bar{x}) = \{f(\bar{x})\}, f$ 在

$\bar{x}$ 自动是 PSNC 的, 则有

$$-\widehat{\partial}^+\varphi_0(\bar{x}) \subset \bigcup \Big[\partial\langle y^*, f\rangle(\bar{x})\Big| y^* \in N(f(\bar{x});\Theta)\Big] + N(\bar{x};\Omega), \tag{5.19}$$

$$0 \in \partial\varphi_0(\bar{x}) + \bigcup \Big[\partial\langle y^*, f\rangle(\bar{x})\Big| y^* \in N(f(\bar{x});\Theta)\Big] + N(\bar{x};\Omega). \tag{5.20}$$

证明 如果 $f\colon X \to Y$ 是 Asplund 空间之间的映射, 且在 $\bar{x}$ 是严格 Lipschitz 的, 则根据定理 3.28 有

$$D_N^* f(\bar{x})(y^*) = \partial\langle y^*, f\rangle(\bar{x}), \quad \forall y^* \in Y^*.$$

于是, 上次微分条件 (5.13) 和下次微分条件 (5.17) 分别简化为 (5.19) 和 (5.20). △

正如前面提到的那样, 上面得到的关于问题 (5.12) 的必要最优性条件是以标准/规范的形式给出的, 该标准形式是由特定的规范条件来保证的. 如果这样的约束规范不成立, 会发生什么? 此时希望得到一广义的非规范形式 (有时称为 Fritz John 形式) 的必要条件, 其中对应于费用函数有一个非负 (可以是 0) 的乘子. 下面给出这种形式的上和下次微分条件, 这实际上由定理 5.7 和定理 5.8 即可导出.

定理 5.11(无约束规范的必要最优性条件) 给定问题 (5.12) 的一个局部最优解 $\bar{x}$, 下述断言成立:

(i) 假设 X 和 Y 是 Banach 空间, $\Omega = X, \Theta = \{0\}, F = f\colon X \to Y$ 在 $\bar{x}$ 是 Fréchet 可微的. 如果 f 在 $\bar{x}$ 是严格可微的且像空间 $\nabla f(\bar{x})X$ 在 Y 中是闭的, 或者 f 在 $\bar{x}$ 附近是连续的且 $\dim Y < \infty$, 那么存在 $\lambda_0 \geqslant 0$, 使得对任意 $x^* \in \widehat{\partial}^+\varphi_0(\bar{x})$, 存在 $y^* \in Y^*$ 满足

$$0 = \lambda_0 x^* + \nabla f(\bar{x})^* y^*, \quad (\lambda_0, y^*) \neq 0. \tag{5.21}$$

(ii) 假设 X 是 Asplund 空间, 而 Y 是 Banach 空间, $f\colon X \to Y$ 在 $\bar{x}$ 是严格可微的, 具有满射导数 $\nabla f(\bar{x})$, 且 Ω 在 $\bar{x}$ 附近是局部闭的. 如果 Ω 在 $\bar{x}$ 是 SNC 的, 或者 Θ 在 $f(\bar{x})$ 是 SNC 的, 那么存在 $\lambda_0 \geqslant 0$, 使得对任意 $x^* \in \widehat{\partial}^+\varphi_0(\bar{x})$, 存在 $y^* \in N(f(\bar{x});\Theta)$ 满足

$$-\lambda_0 x^* - \nabla f(\bar{x})^* y^* \in N(\bar{x};\Omega), \quad (\lambda_0, y^*) \neq 0.$$

(iii) 假设 X 和 Y 都是 Asplund 空间, Ω 和 Θ 是闭的, $S(\cdot) = F(\cdot) \cap \Theta$ 在 $\bar{x}$ 附近是内半紧的. 如果对任意的 $\bar{y} \in S(\bar{x})$, 下列性质之一成立:

(a) Ω 在 $\bar{x}$ 是 SNC 的, 且 F^{-1} 在 $(\bar{y}, \bar{x})$ 是 PSNC 的;

(b) Ω 在 $\bar{x}$ 是 SNC 的, 且 Θ 在 $\bar{y}$ 是 SNC 的;

(c) F 在 $(\bar{x}, \bar{y})$ 是 PSNC 的, 且 Θ 在 $\bar{y}$ 是 SNC 的;

(d) F 在 $(\bar{x}, \bar{y})$ 是 SNC 的.

那么存在 $\lambda_0 \geqslant 0$ 使得对任意的 $x^* \in \partial^+\varphi_0(\bar{x})$, 存在 $\bar{y} \in S(\bar{x})$ 和对偶元 $y^* \in N(\bar{y};\Theta), x_1^* \in D_N^*F(\bar{x},\bar{y})(y^*), x_2^* \in N(\bar{x};\Omega)$ 满足

$$0 = \lambda_0 x^* + x_1^* + x_2^*, \quad (\lambda_0, y^*, x_1^*) \neq 0. \tag{5.22}$$

(iv) 除了 (iii) 中的假设外, 设 φ_0 在 $\bar{x}$ 附近是局部 Lipschitz 的. 如果对任意的 $\bar{y} \in S(\bar{x})$, (iii) 中的性质 (a)~(d) 之一满足, 则存在 $\lambda_0 \geqslant 0, x^* \in \partial\varphi_0(\bar{x}), \bar{y} \in S(\bar{x}), y^* \in N(\bar{y};\Theta), x_1^* \in D_N^*F(\bar{x},\bar{y})(y^*), x_2^* \in N(\bar{x};\Omega)$ 使得 (5.22) 成立.

证明　在所作的假设下, 如果 $\nabla f(\bar{x})\colon X \to Y$ 是满的, 则由定理 5.7(i) 通过取"标准"乘子 $\lambda_0 = 1$ 可得结论 (i) 成立. 如果 $\nabla f(\bar{x})$ 不是满的, 而空间 $\nabla f(\bar{x})X$ 在 Y 中是闭的, 则易证明 (根据分离定理; 请比照定理 1.57 的证明) $\ker \nabla f(\bar{x}^*) \neq \{0\}$, 即存在 $0 \neq y^* \in Y^*$ 使得 $\nabla f(\bar{x})^* y^* = 0$. 因此得 $\lambda_0 = 0, y^* \neq 0$ 时的 (5.21).

下面由定理 5.7(iii) 中的上次微分条件导出 (iii) 中的相应条件. 注意到由定理 5.7 的断言 (ii), (ii) 的证明是完全类似的 (实际上它包含在下面的证明中). 注意到如果规范条件 (5.14) 和 (5.16) 满足, 定理 5.7(iii) 蕴涵着当 $\lambda_0 = 1$ 时想要的结果 (iii). 若 (5.14) 或 (5.16) 不成立, 需证明当 $\lambda_0 = 0$ 时的 (iii) 中的关系成立, 且 $(y^*, x_1^*) \neq 0$. 事实上, 如果 (5.14) 不满足, 则存在 $\bar{y} \in S(\bar{x})$ 和对偶元 $y^* \in N(\bar{y};\Theta), 0 \neq x^* \in D_N^*F(\bar{x},\bar{y})(y^*)$ 使得 $-x^* \in N(\bar{x};\Omega)$. 这就得 (5.22), 其中 $\lambda_0 = 0, x_1^* = x^*, x_2^* = -x^*$. 如果 (5.16) 不满足, 则存在 $\bar{y} \in S(\bar{x}), 0 \neq y^* \in N(\bar{y};\Theta)$ 使得 $0 \in D_N^*F(\bar{x},\bar{y})(y^*)$. 这就得 (5.22), 其中 $\lambda_0 = 0, y^* \neq 0, x_1^* = x_2^* = 0$.

余下的, 需在费用函数 φ_0 在 $\bar{x}$ 附近是 Lipschitz 连续的条件下来证断言 (iv) 中的下次微分必要条件成立. 前面已经提到, 在该条件下 φ_0 在 $\bar{x}$ 自动是 SNEC 的, 并且规范条件 (5.18) 简化为 (5.14). 因此, 在约束规范 (5.14) 和 (5.16) 下, 由定理 5.8 得 (5.22) 对 $\lambda_0 = 1$ 和某个 $x^* \in \partial\varphi_0(\bar{x})$ 成立. 如果 (5.14) 或 (5.16) 不满足, 类似于断言 (iii) 中上次微分条件的证明, 可得到当 $\lambda_0 = 0$ 时 (5.22) 成立.　△

注意到定理 5.11 的断言 (i) 给出了 Banach 空间中具有等式算子约束问题的 Lagrange 乘子法则经典 Lyusternik 版本的一个上次微分推广. 当 f 在 $\bar{x}$ 严格可微时, 简化为本节结果. 当 $\dim Y < \infty, f$ 在 $\bar{x}$ 仅仅 Fréchet 可微时, 这个结果由定理 6.37 和定理 6.38 在等式约束的情形可得; 请比照 6.3.4 小节的证明. 如果规范条件 (5.16) 假定成立且取代定理 5.7 中的 (5.15), 而且如果 φ_0 在定理 5.8(ii) 中假定为局部 Lipschitz 的, 易验证定理 5.11 中的断言 (ii)~(iv) 实际上等价于定理 5.7 和定理 5.8 中相应的断言. 一般来说, 定理 5.7 和定理 5.8 包含更细致的条件保证标准形式的上次微分最优性条件.

有趣的是, 注意到如果 f 在 $\bar{x}$ 仅仅假设为 Fréchet 可微的, 在该点附近关于它没有连续性要求, 则定理 5.11 的断言 (i) 中版本的 Lagrange 乘子法则即使在有限

维空间 X, Y 及线性费用函数 φ_0 的情形也不成立. 这一点由下面的例子来说明.

例 5.12(具有 Fréchet 可微但不连续的等式约束问题乘子法则的反例) 具有 Lagrange 乘子的必要最优性条件可以对极小化线性费用函数的二维问题不成立, 其中的约束是由一个函数给出的等式约束, 而这个函数在全局极小值点处是 Fréchet 可微的, 但在该点附近是不连续的.

证明 考虑约束为

$$0 = f(x_1, x_2) := \begin{cases} x_2^2 + x_1^2, & x_2 \geqslant 0, \\ x_2^2 - x_1^2, & \text{其他} \end{cases}$$

的极小化问题, 其中费用函数 $\varphi_0(x_1, x_2) := x_1$. 易验证 $\bar{x} = (0,0)$ 是该问题的全局极小值点, 这里 f 在 $\bar{x}$ 是 Fréchet 可微的, 但在该点附近不是连续的. 由于 $\nabla\varphi_0(0,0) = (1,0), \nabla f(0,0) = (0,1)$, 所以满足最优性条件 (5.21), 即

$$0 = \lambda_0 \nabla\varphi_0(\bar{x}) + \lambda_1 \nabla f(\bar{x})$$

的唯一数对 (λ_0, λ_1) 只可能是 $(0,0)$, 矛盾. 注意到 f 在 $\bar{x}$ 不是严格可微的. △

在 Asplund 空间之间的映射 $F = f\colon X \to Y$ 是严格 Lipschitz 的情形下, 下面给出定理 5.11 的有效推论.

推论 5.13(无规范的严格 Lipschitz 约束) 设 $\bar{x}$ 是问题 (5.12) 的局部最优解, X 和 Y 是 Asplund 空间, Ω 和 Θ 是闭的, $F = f$ 是单值的且在 $\bar{x}$ 是严格 Lipschitz 的. 如果下列性质之一成立:

(a) Ω 在 $\bar{x}$ 是 SNC 的, 且 f^{-1} 在 $(f(\bar{x}), \bar{x})$ 是 PSNC 的;

(b) Θ 在 $f(\bar{x})$ 是 SNC 的.

则存在 $\lambda_0 \geqslant 0$ 使得对任意的 $x^* \in \widehat{\partial}^+\varphi_0(\bar{x})$, 存在 $y^* \in N(f(\bar{x}); \Theta)$ 满足

$$-\lambda_0 x^* \in \partial\langle y^*, f\rangle(\bar{x}) + N(\bar{x}; \Omega), \quad (\lambda_0, y^*) \neq 0.$$

更进一步, 如果 φ_0 在 $\bar{x}$ 附近是 Lipschitz 连续的, 且 (a) 或 (b) 成立, 则存在 $\lambda_0 \geqslant 0, y^* \in N(f(\bar{x}); \Theta)$ 满足

$$0 \in \lambda_0 \partial\varphi_0(\bar{x}) + \partial\langle y^*, f\rangle(\bar{x}) + N(\bar{x}; \Omega), \quad (\lambda_0, y^*) \neq 0.$$

证明 推论的上和下次微分条件都由定理 5.11 和上导数标量化公式直接可得, 该公式保证在上面的条件下如果 $y^* = 0$, 则 $x_1^* = 0$. 此时定理 5.11 的 (b) 和 (c) 中的条件简化为 Θ 在 $f(\bar{x})$ 的 SNC 性质, 因为由 f 的局部 Lipschitz 连续性可知它在 $\bar{x}$ 自动是 PSNC 的. 下面证明在严格 Lipschitz 映射下定理 5.11(d) 中 f 的

SNC 性质是多余的. 事实上, 根据推论 3.30, 这样的映射 $f\colon X \to Y$ 是 SNC 的当且仅当 Y 是有限维的, 这已包含在关于 Θ 的 SNC 条件中. 因此定理 5.11 的性质 (a)~(d) 简化为推论中的 (a) 和 (b). △

注 5.14(由极点原理得到下次微分条件) 注意到上面所得的下次微分 (而不是上次微分) 必要最优性条件可直接通过极点原理然后由逆像的基本法锥的分析法则和 SNC 性质得到. 事实上, 易见对给定的约束问题 (5.12) 的一个局部最优解 $\bar{x}$, 点 $(\bar{x}, \varphi_0(\bar{x}))$ 是空间 $X \times \mathbb{R}$ 中三个集合

$$\Omega_0 := \text{epi}\,\varphi_0, \quad \Omega_1 := F^{-1}(\Theta) \times \{\varphi_0(\bar{x})\}, \quad \Omega_2 := \Omega \times \{\varphi_0(\bar{x})\}$$

组成的集合系统的局部极点. 把定理 2.22 中的确切极点原理应用于这个系统, 接着像上面那样应用分析法则, 就得到关于 $\bar{x}$ 的用基本法锥和次梯度表示的次微分型必要条件. 注意到, 这种方法不但能得到上面形式的确切/基于点的最优性条件, 而且可得到在无任何 SNC 假设下由局部极小值点附近点的 Fréchet 法向量和次梯度表示的近似/模糊形式的必要条件. 为得到后一形式的必要条件, 需利用定理 2.20 中极点原理的近似版本, 然后利用相应的模糊分析法则; 参见定理 1.14, 引理 3.1 和注 3.21. 本章随后部分将对约束最优化问题 (5.12) 的一些特殊类型和它们相应的多目标最优化问题的变体, 给出这个方向的更多结果.

在这一小节的最后, 考虑由具有无限维值空间的单值映射给出的具有形式 $f(x) = 0$ 的等式型算子约束的特殊类型的最优化问题. 注意到与问题 (5.12) 中的一般约束相比, 上面约束的特殊性是集合 $\Theta = \{0\}$ 从不是 SNC 的, 除非 f 的值空间是有限维的.

对有额外有限多不等式约束以及几何约束的最优化问题, 这里探究了一个富有成效的方法来得到这样问题的必要最优性条件, 即利用某种精确惩罚技术, 将约束问题简化为无约束极小化问题. 下面的相对于几何约束的算子约束映射在参考点处 (而不是像定义 1.47 那样在参考点附近) 的弱化度量正则性性质使这种简化成为可能.

定义 5.15(弱化度量正则性) Banach 空间之间的单值映射 $f\colon X \to Y$ 在点 $\bar{x} \in \Omega \subset X$ 相对于集合 Ω 是度量正则的, 如果存在常数 $\mu > 0, \bar{x}$ 的邻域 U 使得

$$\text{dist}\,(x; S) \leqslant \mu \|f(x) - f(\bar{x})\|, \quad \forall x \in U \cap \Omega,$$

这里 $S := \{x \in \Omega | f(x) = f(\bar{x})\}$.

易见, 如果定义在整个空间 X 上的 Ω- 限制映射 $f_\Omega(x) := f(x) + \Delta(x; \Omega)$ 在定义 1.47(ii) 的意义下在 $\bar{x}$ 附近是局部度量正则的, 则上面的正则性成立. 这样, 第 4 章建立的局部度量正则性的充分条件也保证了在参考点 $\bar{x}$ 处 f 的 Ω- 相对度量正则性成立, 不难发现, 它们对非光滑映射的弱化度量正则性来说一定是不必要的.

这在很大程度上关联于这样的事实：与定义 1.47 的情形相比，定义 5.15 中的度量正则性概念相对于初始点的扰动不是鲁棒的.

下面的结果建立了上面提到类型的约束最优化问题通过某种精确惩罚到无约束问题的所需约化，它方便于在约束极小化下次微分型必要条件上的随后应用.

定理 5.16(等式约束下的精确惩罚) 设 $\bar{x}$ 是约束问题 (CP):

$$\min \quad \varphi_0(x) \quad \text{s.t.} \quad \varphi_i(x) \leqslant 0, \quad i = 1, \cdots, m, \quad f(x) = 0, \quad x \in \Omega$$

的局部最优解，这里 $f\colon X \to Y$ 是 Banach 空间之间的映射，φ_i 是实值函数. 假设 f 在 $\bar{x}$ 附近是局部 Lipschitz 的，在该点相对于 Ω 是度量正则的. 记

$$I(\bar{x}) := \{i \in \{1, \cdots, m\} | \varphi_i(\bar{x}) = 0\},$$

并假设对 $i \in I(\bar{x}) \cup \{0\}$, $\varphi_i(x)$ 在 $\bar{x}$ 附近是局部 Lipschitz 的，对 $i \in \{1, \cdots, m\} \backslash I(\bar{x})$, $\varphi_i(x)$ 在 $\bar{x}$ 是上半连续的. 则对所有充分大的 $\mu > 0$, $\bar{x}$ 是目标函数为

$$\max\left\{\varphi_0(x) - \varphi_0(\bar{x}), \max_{i \in I(\bar{x})} \varphi_i(x)\right\} + \mu\Big(\|f(x)\| + \operatorname{dist}(x; \Omega)\Big)$$

的无约束极小化问题 (UP) 的局部最优解.

证明 易见在对 φ_i 所加的假设下 $\bar{x}$ 是极小化问题

$$\min \varphi(x) := \max\{\varphi_0(x) - \varphi_0(\bar{x}), \max_{i \in I(\bar{x})} \varphi_i(x)\} \quad \text{s.t.} \quad f(x) = 0, \quad x \in \Omega$$

的局部解. 由于 f 在 $\bar{x}$ 是连续的且相对于 Ω 是度量正则的，则存在常数 $\mu_1 > 0$ 和 $\bar{x}$ 的邻域 U 使得对任意 $x \in U \cap \Omega$, 存在 $u \in \Omega$ 满足

$$\varphi(u) \geqslant \varphi(\bar{x}), \quad f(u) = 0, \quad \|x - u\| \leqslant \mu_1 \|f(x)\|.$$

设 ℓ 是 φ 和 f 在 U 上的共同 Lipschitz 常数，且设 $\mu_2 \geqslant \ell\mu_1$. 则对任意的 $x \in U \cap \Omega$, 与 x 相对应的上面的 $u \in \Omega$, 有

$$\begin{aligned}\varphi(x) &\geqslant \varphi(x) - \varphi(u) + \varphi(\bar{x}) \geqslant -\ell\|x - u\| + \varphi(\bar{x}) \\ &\geqslant -\ell\mu_1\|f(x)\| + \varphi(\bar{x}) \geqslant -\mu_2\|f(x)\| + \varphi(\bar{x}),\end{aligned}$$

即 $\bar{x}$ 是问题

$$\min \quad \varphi(x) + \mu_2\|f(x)\| \quad \text{s.t.} \quad x \in \Omega$$

的局部解.

现在注意到 $x \in \Omega$ 等价于 $\operatorname{dist}(x;\Omega)=0$, 该函数在 $\bar{x}$ 相对于 Ω 显然是度量正则的, $\varphi(x)+\mu_2\|f(x)\|$ 是 Lipschitz 的. 应用上面的讨论, 可找到 $\mu_3 \geqslant 0$ 使得 $\bar{x}$ 是问题

$$\min \quad \varphi(x)+\mu_2\|f(x)\|+\mu_3 \operatorname{dist}(x;\Omega)$$

的局部解. 令 $\mu := \max\{\mu_2,\mu_3\}$, 则证明完成. △

基于上面的精确惩罚结果, 并且利用第 3 章的次微分和 SNC 分析法则结果, 以及第 4 章的度量正则性的点基上导数准则, 即可得到定理 5.16 中处理过的 (CP) 型约束问题最优解的有效条件.

定理 5.17(等式型算子约束问题的必要条件) 设 $\bar{x}$ 是问题 (CP) 的局部最优解, X 和 Y 都是 Asplund 空间, 函数 φ_i 满足定理 5.16 的假设, 集合 Ω 在 $\bar{x}$ 附近是局部闭的. 还假设映射 f 在 $\bar{x}$ 是严格 Lipschitz 的, 满足 f_Ω^{-1} 在 $(f(\bar{x}),\bar{x})$ 处是 PSNC 的. 则对 $i \in I(\bar{x}) \cup \{0\}$, 存在实数 $\lambda_i \geqslant 0$, 和线性泛函 $y^* \in Y^*$(这些 λ_i 与 y^* 不同时为零), 满足

$$0 \in \partial\left(\sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i \varphi_i\right)(\bar{x})+\partial\langle g^*, f\rangle(\bar{x})+N(\bar{x};\Omega).$$

证明 首先假设 f 在 $\bar{x}$ 相对于 Ω 是度量正则的, 则存在 $\mu>0$, 使得 $\bar{x}$ 是定理 5.16 中无约束问题 (UP) 的局部最优解. 因此

$$0 \in \partial\Big(\max\big\{\varphi_0(\cdot)-\varphi_0(\bar{x}), \max_{i \in I(\bar{x})} \varphi_i(\cdot)\big\}+\mu(\|f(\cdot)\|+\operatorname{dist}(\cdot;\Omega))\Big)(\bar{x}).$$

现对上面的函数应用定理 3.36 中的次微分和法则, 然后应用定理 3.46(ii) 中的极大值函数法则, 对复合 $\|f(x)\|=(\psi \circ f)(x)$, 其中 $\psi(y):=\|y\|$, 应用推论 3.43 中的链式法则, 对定理 1.97 中的距离函数 $\operatorname{dist}(x;\Omega)$ 应用次微分公式, 则得满足 $(\lambda_i | i \in I(\bar{x})) \cup \{0\}) \neq 0$ 的定理的必要最优性条件.

如果不假设 f 在 $\bar{x}$ 相对 Ω 是度量正则的, 则映射 $f_\Omega(x):=f(x)+\Delta(x;\Omega)$ 在定义 1.47(ii) 的意义下在 $\bar{x}$ 附近不是度量正则的. 根据定理 4.18(c), 有 $\ker \widetilde{D}_M^* f_\Omega(\bar{x}) \neq \{0\}$ 或 f_Ω^{-1} 在 $(f(\bar{x}),\bar{x})$ 不是 PSNC 的. 根据这个定理的假设后者是不可能的. 因此存在 $y^* \neq 0$ 使得

$$0 \in \widetilde{D}_M^* f_\Omega(\bar{x})(y^*) \subset D_N^* f_\Omega(\bar{x})(y^*)=D_N^*\Big(f+\Delta(\cdot;\Omega)\Big)(\bar{x})(y^*).$$

应用命题 3.12 中上导数和法则 (由 f 的 Lipschitz 连续性命题 3.12 的规范假设成立), 再利用定理 3.28 中的标量化公式, 由 f 是严格 Lipschitz 的, 得包含关系

$$0 \in D_N^* f(\bar{x})(y^*)+N(\bar{x};\Omega)=\partial\langle y^*, f\rangle(\bar{x})+N(\bar{x};\Omega).$$

这保证满足 $y^* \neq 0$ 的定理的结论成立. △

注意到如果 f 假设为在 $\bar{x}$ 附近仅仅是 Lipschitz 连续的 (而非在该点是严格 Lipschitz 的), 则定理 5.17 的结论成立且具有形式

$$0 \in \partial\left(\sum_{i \in I(\bar{x}) \cup \{0\}} \lambda_i \varphi_i\right)(\bar{x}) + D_N^* f(\bar{x})(y^*) + N(\bar{x}; \Omega),$$

其中 $(\lambda_i, y^*) \neq 0$. 这由定理的证明直接可得.

下面的推论描述了涉及广义 Fredholm 映射的一大类算子约束, 它们满足上面定理的假设. 这个结果对在最优控制问题中的应用尤为重要, 见第 6 章.

推论 5.18(具有广义 Fredholm 算子约束的最优化问题的必要条件) 设 $\bar{x}$ 是具有算子约束的上面问题 (CP) 的局部最优解. 假设 f 在 $\bar{x}$ 是广义 Fredholm 的, Ω 在 $\bar{x}$ 是 SNC 的, 而且 (CP) 中所有其他数据满足定理 5.17 的假设. 则该定理的必要最优性条件成立.

证明 正如定理 3.35 的证明那样, 在对 f 和 Ω 所加的假设下, f_Ω^{-1} 在 $(f(\bar{x}), \bar{x})$ 是 PSNC 的. 由于每个紧严格 Lipschitz 映射自动是严格 Lipschitz 的, 而且加上一个线性有界算子不破坏这个性质, 从而 f 在 $\bar{x}$ 是严格 Lipschitz 的, 这样就完成了推论的证明. △

5.1.3 泛函约束下的必要条件

这一小节更详细地研究一类特殊的约束问题 (5.12), 其约束是由有限多无限维空间上的实值函数定义的等式型和不等式型泛函约束. 即给定 $\varphi_i\colon X \to \mathbb{R}, i = 0, 1, \cdots, m+r, \Omega \subset X$, 考虑下面的非可微规划问题:

$$\begin{cases} \min & \varphi_0(x) \\ \text{s.t.} & \varphi_i(x) \leqslant 0, \quad i = 1, \cdots, m, \\ & \varphi_i(x) = 0, \quad i = m+1, \cdots, m+r, \\ & x \in \Omega. \end{cases} \tag{5.23}$$

注意到函数 φ_i 其实可以是增广实值的. 由于 (5.23) 中有额外的几何约束, 故这里的实值性假设并没限制问题的一般性. 显然 (5.23) 是当 $F = (\varphi_1, \cdots, \varphi_{m+r})\colon X \to \mathbb{R}^{m+r}$ 和

$$\begin{aligned} \Theta = &\{(\alpha_1, \cdots, \alpha_{m+r}) \in \mathbb{R}^{m+r} | \alpha_i \leqslant 0, \quad i = 1, \cdots, m, \\ &\text{且}\, \alpha_i = 0, \quad i = m+1, \cdots, m+r\} \end{aligned} \tag{5.24}$$

时 (5.12) 的特殊情形. 因此由 (5.24) 中集合 Θ 的形式, 5.1.2 小节的结果直接蕴涵着问题 (5.23) 的必要最优性条件. 然而, (5.23) 的特殊结构还可以得到关于局部极小值比由 (5.12) 产生的必要条件更精细的结果.

下面首先建立 (5.23) 的局部极小值的上次微分条件. 下面的定理含特别针对不等式约束的新结果, 也包括源于 5.1.2 小节中结果的对问题 (5.23) 的必要最优条件. 和往常一样, 对在有限维空间中取值的映射的基本上导数使用相同的上导数记号 D^*. 简单起见, 这里只给出无约束规范的必要最优性条件; 标准形式 (即有规范条件时) 的这些条件或者来自 5.1.2 小节中的相应结果, 或者用类似的方法可得.

定理 5.19(非可微规划的上次微分条件)　设 $\bar{x}$ 是问题 (5.23) 的局部最优解, 集合 Ω 在 $\bar{x}$ 附近是局部闭的, 对 $i = m+1, \cdots, m+r$, 函数 φ_i 在该点附近是连续的. 则下面的断言成立:

(i) 假设 X 有一个 Lipschitz $\mathcal{C}^1$ 阻尼函数 (当 X 有一个 Fréchet 可微重赋范时, 这是自动成立的, 特别地, 当 X 是自反的), 假设 Ω 或 $f := (\varphi_{m+1}, \cdots, \varphi_{m+r})$ 在 $\bar{x}$ 是 SNC 的. 则对任意 Fréchet 上次梯度 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x}),\quad i = 0, \cdots, m$, 存在 $(\lambda_0, \cdots, \lambda_{m+r}) \in \mathbb{R}^{m+r+1}, x^* \in D^*f(\bar{x})(\lambda_{m+1}, \cdots, \lambda_{m+r})$, 和 $\widetilde{x}^* \in N(\bar{x}; \Omega)$ 满足关系:

$$\lambda_i \geqslant 0 \quad (i = 0, \cdots, m), \quad \lambda_i\varphi_i(\bar{x}) = 0 \quad (i = 1, \cdots, m), \tag{5.25}$$

$$0 = \sum_{i=0}^{m} \lambda_i x_i^* + x^* + \widetilde{x}^*, \quad (\lambda_0, \cdots, \lambda_{m+r}, x^*) \neq 0. \tag{5.26}$$

如果 φ_i 在 $\bar{x}$ 附近是 Lipschitz 连续的 $(i = m+1, \cdots, m+r)$, 则除了 (5.25) 外还有

$$-\sum_{i=0}^{m} \lambda_i x_i^* \in \partial\left(\sum_{i=m+1}^{m+r} \lambda_i\varphi_i\right)(\bar{x}) + N(\bar{x}; \Omega), (\lambda_0, \cdots, \lambda_{m+r}) \neq 0, \tag{5.27}$$

其中仅假设 Ω 是局部闭的, 不需要 (φ_i, Ω) 上的其他条件.

(ii) 假设 X 是 Asplund 空间, $f := (\varphi_1, \cdots, \varphi_{m+r})$ 在 $\bar{x}$ 附近是连续的, Ω 或 f 在 $\bar{x}$ 是 SNC 的. 则存在 $\lambda_0 \geqslant 0$, 使得对任意 Fréchet 上次梯度 $x_0^* \in \widehat{\partial}^+\varphi_0(\bar{x})$, 存在 $(\lambda_1, \cdots, \lambda_{m+r}) \in \mathbb{R}^{m+r}, x^* \in D^*f(\bar{x})(\lambda_1, \cdots, \lambda_{m+r})$ 和 $\widetilde{x}^* \in N(\bar{x}; \Omega)$ 满足 (5.25) 和

$$0 = \lambda_0 x_0^* + x^* + \widetilde{x}^*, \quad (\lambda_0, \cdots, \lambda_{m+r}, x^*) \neq 0. \tag{5.28}$$

如果对 $i = 1, \cdots, m+r$, φ_i 在 $\bar{x}$ 附近是 Lipschitz 连续的, 则除了 (5.25) 外, 还有

$$-\lambda_0 x_0^* \in \partial\left(\sum_{i=1}^{m+r} \lambda_i\varphi_i\right)(\bar{x}) + N(\bar{x}; \Omega), \quad (\lambda_0, \cdots, \lambda_{m+r}) \neq 0, \tag{5.29}$$

其中仅设 Ω 是局部闭的, 不需要 (φ_i, Ω) 上的其他条件.

证明　为在一般的假设下证 (i) 成立, 对 $i = 0, \cdots, m$, 任取 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$, 并且把当 $\mathcal{S} = \mathcal{LC}^1$ 时的定理 1.88(ii) 中的变分描述应用于次梯度 $-x_i^* \in \widehat{\partial}(-\varphi_i)(\bar{x})$. 由此可找到函数 $s_i \colon X \to \mathbb{R},\ i = 0, \cdots, m$ 满足

$$s_i(\bar{x}) = \varphi_i(\bar{x}), \quad 在\ \bar{x}\ 附近\ \ s_i(x) \geqslant \varphi_i(x),$$

每个 $s_i(x)$ 在 $\bar{x}$ 附近是连续可微的, 且 $\nabla s_i(\bar{x}) = x_i^*$. 易验证 $\bar{x}$ 是下面 (5.23) 型的最优化问题的局部解, 其中费用函数和不等式约束函数在 $\bar{x}$ 附近是连续可微的:

$$\begin{cases} \min & s_0(x) \\ \text{s.t.} & s_i(x) \leqslant 0, \quad i = 1, \cdots, m, \\ & \varphi_i(x) = 0, \quad i = m+1, \cdots, m+r, \\ & x \in \Omega. \end{cases} \tag{5.30}$$

现对问题 (5.30) 应用定理 5.11(iii) 中的必要条件. 该问题对应于 (5.12) 的 $F := (s_1, \cdots, s_m, \varphi_{m+1}, \cdots, \varphi_{m+r})$ 和集合 Θ 由 (5.24) 定义的情形. 注意到

$$\begin{aligned} &N\Big((\varphi_1(\bar{x}), \cdots, \varphi_{m+r}(\bar{x})); \Theta\Big) \\ =&\big\{(\lambda_1, \cdots, \lambda_{m+r}) \in \mathbb{R}^{m+r} | \lambda_i \geqslant 0, \lambda_i \varphi_i(\bar{x}) = 0, \quad i = 1, \cdots, m\big\}, \end{aligned}$$

其中 $s_i(\bar{x}) = \varphi_i(\bar{x}), i = 1, \cdots, m$, 并注意到对上面的 F 有

$$F(x) = (s(x), 0) + \Big(0, \varphi_{m+1}(x), \cdots, \varphi_{m+r}(x)\Big), \tag{5.31}$$

其中 $s := (s_1, \cdots, s_m)\colon X \to \mathbb{R}^m$ 在 $\bar{x}$ 附近是连续可微的. 因此对 $y^* = (\lambda_1, \cdots, \lambda_{m+r})$, 定理 5.11(iii) 中的条件 $y^* \in N(\bar{y}; \Theta)$ 简化为当 $i = 1, \cdots, m$ 时 (5.25) 中的符号和互补松弛条件.

由于在定理 5.11(iii) 中 $Y = \mathbb{R}^{m+r}$, 所以根据定理 1.70, (5.31) 中 F 的 SNC 和 PSNC 性质等价于 $f = (\varphi_{m+1}, \cdots, \varphi_{m+r})$ 的 SNC 性质, 易见定理 5.11(iii) 中的条件 (a)~(d) 之一成立当且仅当 Ω 或 f 在 $\bar{x}$ 是 SNC 的. 对 (5.31) 中的和式应用定理 1.62(ii) 中的上导数和法则可知, 在 $x_1^* \in D^*F(\bar{x}, \bar{y})(y^*), x_2^* \in N(\bar{x}; \Omega)$ 时的 (5.22) 等价于

$$0 = \sum_{i=0}^{m} \lambda_i \nabla s_i(\bar{x}) + x^* + \widetilde{x}^*, \quad (\lambda_0, \cdots, \lambda_{m+r}, \widetilde{x}^*) \neq 0,$$

其中 $x^* \in D^*f(\bar{x})(\lambda_{m+1}, \cdots, \lambda_{m+r})$, $\widetilde{x}^* \in N(\bar{x}; \Omega)$ 和 $\lambda_0 \geqslant 0$. 考虑到对 $i = 0, \cdots, m, \nabla s_i(\bar{x}) = x_i^*$, 则得 (5.26). 由 (5.26) 在 φ_i $(i = m+1, \cdots, m+r)$ 是局部 Lipschitz 时得 (5.27), 只需注意到此时 f 在 $\bar{x}$ 自动是 SNC 的, 然后对上导数 $D^*f(\bar{x})$ 应用标量化公式, 即有

$$D^*f(\bar{x})(\lambda_{m+1}, \cdots, \lambda_{m+r}) = \partial\left(\sum_{i=m+1}^{m+r} \lambda_i \varphi_i\right)(\bar{x}).$$

下面来证 (ii). 为此取 $F = f := (\varphi_1, \cdots, \varphi_{m+r})$, Θ 与 (5.24) 中定义的相同, 然后直接应用定理 5.11(iii). 用这种方法, 在 (ii) 的一般假设下对某 $x^* \in D^*f(\bar{x})(\lambda_1, \cdots,$

λ_{m+r}) 有 (5.25) 和 (5.28) 成立. 当所有的 $\varphi_1, \cdots, \varphi_{m+r}$ 在 $\bar{x}$ 附近是 Lipschitz 连续的, 根据上导数标量化 (5.25) 和 (5.28) 得 (5.29). △

值得注意的是定理 5.19 中的必要条件或者由"集成"映射 $(\varphi_{m+1}, \cdots, \varphi_{m+r})\colon X \to \mathbb{R}^r$ 和 $(\varphi_1, \cdots, \varphi_{m+r})\colon X \to \mathbb{R}^{m+r}$ 的上导数给出, 或者通过 (5.27) 和 (5.29) 中和的次梯度给出. 基于上导数和次微分分析法则, 这些条件可以用单值函数 φ_i 的上导数和次梯度的分离的形式来表达, 只是结果要弱一点. 特别地, 对定理 5.19 中的上导数结果, 可应用定理 3.10 中的上导数和法则于

$$f(x) = \Big(\varphi_{m+1}(x), 0, \cdots, 0\Big) + \cdots + \Big(0, \cdots, 0, \varphi_{m+r}(x)\Big),$$

然后利用定理 1.80 和定理 2.40, 由 φ_i 和 $-\varphi_i$ 的基本和奇异次梯度来表示 φ_i 的上导数. 简略起见, 下面只对 Lipschitz 函数 φ_i 给出这种形式的结果, 此时相应的条件由定理 3.36 的次微分分析法则可简单地得出. 在这种情形中, 对每个函数 $\varphi_i, i = m+1, \cdots, m+r$, 应用双侧对称次微分

$$\partial^0\varphi(\bar{x}) := \partial\varphi(\bar{x}) \cup \partial^+\varphi(\bar{x})$$

来描述所论的优化问题 (5.23) 中的等式约束是方便的.

推论 5.20(等式约束的具有对称次微分的上次微分条件) 设 $\bar{x}$ 是问题 (5.23) 的局部最优解, 集合 Ω 在 $\bar{x}$ 附近是局部闭的, 对 $i = m+1, \cdots, m+r$, 函数 φ_i 在该点附近是 Lipschitz 连续的. 则下面的断言成立:

(i) 假设 X 有一个 Lipschitz $\mathcal{C}^1$ 阻尼函数. 则对任意的 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x}), i = 0, \cdots, m$, 存在乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$ 满足 (5.25) 和

$$-\sum_{i=0}^{m} \lambda_i x_i^* \in \sum_{i=m+1}^{m+r} \lambda_i \partial^0\varphi_i(\bar{x}) + N(\bar{x}; \Omega).$$

(ii) 假设 X 是 Asplund 空间, 对 $i = 1, \cdots, m$, φ_i 在 $\bar{x}$ 附近也是 Lipschitz 连续的. 则存在 $\lambda_0 \geqslant 0$, 使得对任意的 Fréchet 上次梯度 $x_0^* \in \widehat{\partial}^+\varphi_0(\bar{x})$, 存在乘子 $(\lambda_1, \cdots, \lambda_{m+r}) \in \mathbb{R}^{m+r}$ 满足 (5.25) 和

$$-\lambda_0 x_0^* \in \sum_{i=1}^{m} \lambda_i \partial\varphi_i(\bar{x}) + \sum_{i=m+1}^{m+r} \lambda_i \partial^0\varphi_i(\bar{x}) + N(\bar{x}; \Omega), \quad (\lambda_0, \cdots, \lambda_{m+r}) \neq 0.$$

证明 (i) 中的包含关系根据定理 3.36 的次微分和法则、关系 $\partial(\lambda\varphi)(\bar{x}) = \lambda\partial\varphi(\bar{x})$ (当$\lambda \geqslant 0$) 和 $\partial(\lambda\varphi)(\bar{x}) \subset \lambda\partial^0\varphi(\bar{x})$ (当$\lambda \in \mathbb{R}$), 由 (5.27) 可得. 类似地, 可由定理 5.19(ii) 的 (5.29) 得 (ii) 中的包含关系. △

利用单个函数来描述具有等式约束的问题的必要最优性条件的另一方法 (实际上比推论 5.20 中的方法更精细) 是利用由

$$\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})$$

给出的 φ 在 $\bar{x}$ 的“偶次微分”集, 此时只有非负的乘子. 这是由于

$$\partial(\lambda\varphi)(\bar{x}) \subset |\lambda|\big[\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})\big], \quad \forall\lambda \in \mathbb{R}. \tag{5.32}$$

下面将应用这个描述. 注意到上面的“偶次微分”集合对函数 φ 和 $-\varphi$ 来说相同; 这也是其中名称的由来, 尽管当 φ 是光滑的时, 集合 $\partial\varphi(\bar{x}) \cup \partial(-\varphi)(\bar{x})$ 不简化为经典的梯度.

下面建立具有不等式、等式和几何约束的问题 (5.23) 的下次微分型必要最优性条件, 即与 5.1.2 小节类似, 这样类型的必要最优性条件涉及费用和不等式约束函数的下次梯度或它们上图的法向量, 而不是它们的上次梯度. 根据初始数据的不同假设, 这里在此方向上建立数个结果, 所用的技术也是各不相同的. 与上次微分的情形一样, 这里把焦点放在无约束规范的一般最优性条件, 其与标准 (规范) 形式的关系与上一节的情形是一样的.

这种形式的第一个定理给出的问题 (5.23) 的必要最优性条件是由分离的约束的法向量和次梯度描述的. 这基于极点原理的直接应用, 甚至没有应用任何分析法则. 根据在证明中使用的极点原理的相应版本, 将给出近似和确切形式的必要条件. 对具有 Lipschitz 数据的问题则给出了更具体的确切条件.

定理 5.21(由分离的约束的法向量和次梯度表示的必要条件) 设 $\bar{x}$ 是问题 (5.23) 的局部最优解, 空间 X 是 Asplund 空间, 集合 Ω 在 $\bar{x}$ 附近是局部闭的. 则下面的断言成立:

(i) 假设函数 φ_i 在 $\bar{x}$ 附近是 l.s.c.的 $(i = 0, \cdots, m)$, 在该点附近是连续的 $(i = m+1, \cdots, m+r)$. 则对任意的 $\varepsilon > 0$, 存在点

$$(x_0, \alpha_0) \in \operatorname{epi}\varphi_0 \cap \Big[\big(\bar{x}, \varphi_0(\bar{x})\big) + \varepsilon\mathbb{B}\Big], \quad \widehat{x} \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B}),$$

$$(x_i, \alpha_i) \in \operatorname{epi}\varphi_i \cap \Big[(\bar{x}, 0) + \varepsilon\mathbb{B}\Big], \quad i = 1, \cdots, m,$$

$$(x_i, \alpha_i) \in \operatorname{gph}\varphi_i \cap \Big[(\bar{x}, 0) + \varepsilon\mathbb{B}\Big], \quad i = m+1, \cdots, m+r$$

和对偶元

$$(x_i^*, -\lambda_i) \in \widehat{N}((x_i, \alpha_i); \operatorname{epi}\varphi_i) + \varepsilon\mathbb{B}^*, \quad i = 0, \cdots, m,$$

$$(x_i^*, -\lambda_i) \in \widehat{N}\big((x_i, \alpha_i); \operatorname{gph}\varphi_i\big) + \varepsilon\mathbb{B}^*, \quad i = m+1, \cdots, m+r,$$

$$\widehat{x}^* \in \widehat{N}(\widehat{x}; \Omega) + \varepsilon\mathbb{B}^*$$

满足关系

$$x_0^* + \cdots + x_{m+r}^* + \widehat{x}^* = 0, \tag{5.33}$$

$$\|(x_0^*, \lambda_0)\| + \cdots + \|(x_{m+r}^*, \lambda_{m+r})\| + \|\widehat{x}^*\| = 1. \tag{5.34}$$

(ii) 假设除其中一个集合外所有集合 $\operatorname{epi}\varphi_i$ $(i=0,\cdots,m)$, $\operatorname{ghp}\varphi_i$ $(i=m+1,\cdots,m+r)$ 和 Ω 分别在点 $(\bar{x},\varphi_0(\bar{x})),(\bar{x},0)$ 和 $\bar{x}$ 处是 SNC 的. 则存在

$$(x_0^*,-\lambda_0)\in N\Big((\bar{x},\varphi_0(\bar{x}));\operatorname{epi}\varphi_0\Big),\quad \widehat{x}^*\in N(\widehat{x};\Omega),$$

$$(x_i^*,-\lambda_i)\in N\big((\bar{x},0);\operatorname{epi}\varphi_i\big),\quad i=1,\cdots,m,$$

$$(x_i^*,-\lambda_i)\in N\big((\bar{x},0);\operatorname{ghp}\varphi_i\big),\quad i=m+1,\cdots,m+r$$

满足关系 (5.33) 和 (5.34), 其中 $\lambda_i\geqslant 0$, $i=0,\cdots,m$. 更进一步, 如果对那些满足 $\varphi_i(\bar{x})<0$ 的 $i=1,\cdots,m$, φ_i 还被假设为在 $\bar{x}$ 是上半连续的, 则

$$\lambda_i\varphi_i(\bar{x})=0,\quad i=1,\cdots,m.$$

(iii) 假设对所有的 $i=0,\cdots,m+r$, 函数 φ_i 在 $\bar{x}$ 附近是 Lipschitz 连续的. 则存在乘子 $(\lambda_0,\cdots,\lambda_{m+r})\neq 0$, 使得

$$0\in\sum_{i=0}^{m}\lambda_i\partial\varphi_i(\bar{x})+\sum_{i=m+1}^{m+r}\lambda_i\Big[\partial\varphi_i(\bar{x})\cup\partial(-\varphi_i)(\bar{x})\Big]+N(\bar{x};\Omega),$$

$\lambda_i\geqslant 0$ 对所有的 $i=0,\cdots,m+r$ 成立, 并且对 $i=1,\cdots,m$ 有 $\lambda_i\varphi_i(\bar{x})=0$.

证明 为证 (i), 不失一般性, 假设 $\varphi_0(\bar{x})=0$. 于是易见 $(\bar{x},0)$ 是下面 Apslund 空间 $X\times\mathbb{R}^{m+r+1}$ 中的闭集系统:

$$\Omega_i:=\big\{(x,\alpha_0,\cdots,\alpha_{m+r})\,\big|\,\alpha_i\geqslant\varphi_i(x)\big\},\quad i=0,\cdots,m,$$

$$\Omega_i:=\big\{(x,\alpha_0,\cdots,\alpha_{m+r})\,\big|\,\alpha_i=\varphi_i(x)\big\},\quad i=m+1,\cdots,m+r,$$

$$\Omega_{m+r+1}:=\Omega\times\{0\}$$

的局部极点. 则 (i) 中的近似最优性条件由定理 2.20 中的极点原理的近似版本直接可得. 类似地, 利用定理 2.22 中在 SNC 假设下极点原理的确切版本, 找到元素 (x^*,λ_i) 和 $\widehat{x}^*$ 满足 (5.33), (5.34) 和 (ii) 中的法锥包含关系. 由关于上图的基本法向量的命题 1.76 得 $\lambda_i\geqslant 0$ $(i=0,\cdots,m)$. 为建立 (ii) , 余下需证明在关于 φ_i 额外的假设下互补松弛条件成立. 事实上, 若对某个 $i\in\{1,\cdots,m\}$ 有 $\varphi_i(\bar{x})<0$, 则如果 φ_i 在 $\bar{x}$ 是上半连续的, 对所有 $\bar{x}$ 附近的 x 有 $\varphi_i(x)<0$. 这蕴涵着 $(\bar{x},0)$ 是 φ_i 的上图的内点. 因此对这个 i, 有 $N\big((\bar{x},0);\operatorname{epi}\varphi_i\big)=\{0\}$ 和 $x_i^*=\lambda_i=0$, 这就完成了断言 (ii) 的证明.

为证 (iii), 注意到, 如果 φ 在 $\bar{x}$ 附近是 Lipschitz 连续的, 则根据命题 1.76 和推论 1.81 有

$$(x^*,-\lambda)\in N\Big((\bar{x},\varphi(\bar{x}));\operatorname{epi}\varphi\Big)\Leftrightarrow x^*\in\lambda\partial\varphi(\bar{x}),\quad \lambda\geqslant 0.$$

另一方面, 根据局部 Lipschitz 函数的上导数标量化, 有

$$(x^*, -\lambda) \in N\Big((\bar{x}, \varphi(\bar{x})); \text{ghp}\varphi\Big) \Leftrightarrow x^* \in D^*\varphi(\bar{x})(\lambda) = \partial\langle\lambda, \varphi\rangle(\bar{x}).$$

最后利用 (5.32), 同时考虑到 (5.33) 和 (5.34), 就完成了 (iii) 及整个定理的证明. △

注 5.22(不同形式的必要最优性条件的比较) 类似推论 5.20, 对等式约束的情形, 可以写出更通常形式的必要最优性条件, 即由具有可为任意实数的乘子 λ_i 的双侧次微分 $\partial^0\varphi_i(\bar{x})$ 取代具有非负乘子 λ_i 的偶次微分 $\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})$. 由定义知, 偶次微分形式的必要条件比对称次微分更精细. 下面通过在 $\mathbb{R}^2$ 中的例子来说明这个事实:

$$\min \quad x_1 \quad \text{s.t.} \quad \varphi(x_1, x_2) := \big||x_1| + x_2\big| + x_1 = 0.$$

基于例 2.49 中对函数 φ 的次梯度的计算, 根据定理 5.21(iii), 得 $\bar{x} = (0,0)$ 不是上面问题的最优解, 而当 $\lambda \in \mathbb{R}$ 时由 $\lambda\partial^0\varphi(0)$ 却得不出这样的结论. 当然, 通过利用 φ 的 Clarke 广义梯度 $\partial_C\varphi(\bar{x})$ 和 Warga 最小导容 $\Lambda^0\varphi(\bar{x})$ 也得不出这个结论, 因为这两个双侧次微分结构总是包含 $\partial^0\varphi(\bar{x})$ 的, 通过例 2.49 的计算可说明这一点.

由于 $\partial_C\varphi(\bar{x})$ 和 $\Lambda^0\varphi(\bar{x})$ 都可以本质地比 $\partial\varphi(\bar{x})$ 大 (从来没有更小), 上面结果中基本次微分的使用就得到了关于具有非光滑费用函数和不等式约束问题的局部极小值的更精细的必要最优性条件. 最简单的用来说明的例子是下面无约束的一维问题

$$\min \quad \varphi(x) := -|x|, \quad x \in \mathbb{R},$$

其中 $\bar{x} = 0$ 不是极小值 (而是极大值) 点, 而 $0 \in \partial_C\varphi(0) = [-1,1]$. 另一方面, $0 \notin \partial\varphi(0) = \{-1, 1\}$.

对二维问题

$$\min \quad x_1 \quad \text{s.t.} \quad \varphi(x_1, x_2) := |x_1| - |x_2| \leqslant 0,$$

有 $\partial\varphi(0,0) = \{(v_1, v_2) | -1 \leqslant v_1 \leqslant 1, v_2 = 1 \ \text{或} \ v_2 = -1\}$, 因此, 根据定理 5.21(iii), 点 $\bar{x} = (0,0)$ 可从最优解中排除, 而 $\partial_C\varphi(0,0) = \{(v_1, v_2) | -1 \leqslant v_1 \leqslant 1, -1 \leqslant v_2 \leqslant 1\}$ 却做不到这一点. 具有非光滑不等式约束的二维问题的另一例子是

$$\min \quad x_2 \quad \text{s.t.} \quad \varphi(x_1, x_2) := \big||x_1| + x_2\big| + x_2 \leqslant 0,$$

其中 $\partial\varphi(0,0) = \{(v_1, v_2) \,\big|\, |v_1| + 1 \leqslant v_2 \leqslant 2\} \cup \{(v_1, v_2) \,\big|\, 0 \leqslant v_2 \leqslant -|v_1| + 1\}$; 更多细节参见例 2.49. 这样定理 5.21(iii) 的结果就可以排除掉非最优点 $\bar{x} = (0,0)$, 而利用 $\partial_C\varphi(0,0)$ 或 $\Lambda^0\varphi(0,0)$ 则不能.

下面要建立的最优化结果具有 Lagrange 原理的形式, 也就是说, 约束问题的必要最优性条件可由用合适乘子基于原始约束的一些 Lagrange 函数的无约束局部

极小值的必要条件给出. 对极小化问题 (5.23), 这里考虑涉及费用函数和泛函 (不是几何) 约束的标准 Lagrange 函数:

$$L(x,\lambda_0,\cdots,\lambda_{m+r}):=\lambda_0\varphi_0(x)+\cdots+\lambda_{m+r}\varphi_{m+r}(x), \tag{5.35}$$

也考虑涉及几何约束的“本质 Lagrange 函数”:

$$L_\Omega(x;\lambda_0,\cdots,\lambda_{m+r}):=\lambda_0\varphi_0(x)+\cdots+\lambda_{m+r}\varphi_{m+r}(x)+\delta(x;\Omega). \tag{5.36}$$

为得到 Lagrange 准则型的一般结果, 首先建立一个关于分析法则的引理, 这个引理无疑具有独立的意义, 并且在后面的结果中也将用到. 给定 Banach 空间之间的单值映射 $f\colon X\to Z$, 子集 $\Omega\subset X$ 和 $\Theta\subset Z$, 考虑集合

$$\mathcal{E}(f,\Omega,\Theta):=\{(x,z)\in X\times Z\,\big|\,f(x)-z\in\Theta,x\in\Omega\}, \tag{5.37}$$

它可以看作函数 f 在 Ω 上相对于 Θ 的广义上图. 特别地, 如果 $f=(\varphi_0,\cdots,\varphi_m)\colon X\to\mathbb{R}^{m+1}$, $\Theta=\mathbb{R}^{m+1}_-$ 是 $Z=\mathbb{R}^{m+1}$ 的非正象限, 则集合 (5.37) 是向量函数 f 相对于 $\mathbb{R}^{m+1}$ 上的标准序的上图. 对 $\Theta=\{0\}$, 集合 (5.37) 刚好是 f 的图像. 更一般地, 如果 $\Theta\subset Z$ 是为 Z 定序的凸锥, 则 (5.37) 是映射 $f\colon X\to Z$ 在集合 Ω 上的限制 $f_\Omega:=f|_\Omega$ 相对于 Z 上的这个序的上图. 注意到总可以通过指示映射 $\Delta(\cdot;\Omega)$ ($\Delta(x;\Omega):=0\in Z,$若$x\in\Omega$ 和$\Delta(x):=\varnothing,$其他) 来使用符号

$$f_\Omega(x)=f(x)+\Delta(x;\Omega),\quad\forall x\in X.$$

在下面的引理中利用定义 4.8 中给出的强上导数的正规性性质; 这个性质的一些充分条件见命题 4.9.

引理 5.23(广义上图的基本法向量) 设 $f\colon X\to Z$ 是 Banach 空间之间的映射, $\Omega\subset X,\Theta\subset Z$ 满足 $\bar{x}\in\Omega,f(\bar{x})-\bar{z}\in\Theta$. 则下面的断言成立:

(i) 假设 f 在 $\bar{x}$ 附近相对于 Ω 是局部 Lipschitz 的, 则

$$D^*_Mf_\Omega(\bar{x})(z^*)=\partial\langle z^*,f_\Omega\rangle(\bar{x}),\quad\forall z^*\in Z^*.$$

(ii) 如果 f 在 $\bar{x}$ 相对于 Ω 是连续的, 则

$$(x^*,z^*)\in N((\bar{x},\bar{z});\mathcal{E}(f,\Omega,\Theta))\Rightarrow -z^*\in N(f(\bar{x})-\bar{z};\Theta).$$

进一步假设 X 和 Z 都是 Asplund 空间, Ω 和 Θ 分别在 $\bar{x}$ 和 $f(\bar{x})-\bar{z}$ 附近是局部闭的, 则

$$\begin{aligned}N((\bar{x},\bar{z});\mathcal{E}(f,\Omega,\Theta))\subset\{(x^*,z^*)\in X^*\times Z^*\,\big|\,&x^*\in D^*_Nf_\Omega(\bar{x})(z^*),\\&-z^*\in N(f(\bar{x})-\bar{z};\Theta)\}.\end{aligned} \tag{5.38}$$

(iii) 除了 (ii) 中的假设外, 还假设 f 在 $\bar{x}$ 附近相对于 Ω 是局部 Lipschitz 的, f_Ω 在 $\bar{x}$ 是强上导数正规的, 则

$$N\big((\bar{x},\bar{z});\mathcal{E}(f,\Omega,\Theta)\big)\subset\big\{(x^*,z^*)\in X^*\times Z^*\,\big|\,x^*\in\partial\langle z^*,f_\Omega\rangle(\bar{x}),\\ -z^*\in N(f(\bar{x})-\bar{z};\Theta)\big\}.\tag{5.39}$$

(iv) 假设 f 在 $\bar{x}$ 相对于 Ω 是 w^* 严格 Lipschitz 的, 则在任意 Banach 空间 X 和 Z 的情形下 (5.39) 中相反的包含关系成立. 更进一步, 如果 f_Ω 在 $\bar{x}$ 是强上导数正规的, 则 (5.38) 中相反的包含关系也成立.

证明 断言 (i) 是定理 1.90 中混合标量化公式的推广, 可用完全相同的方法证明. 实际上, 与引理 3.27 和定理 3.28 中基本 (法锥) 标量化公式的证明相比, 定理 1.90 的证明中从没应用 $\Omega=X$ 的线性结构. 注意到断言 (i) 给断言 (ii) 和 (iii) 架起了一座桥梁.

根据 (5.37) 中集合 $\mathcal{E}(f,\Omega,\Theta)$ 的结构, (ii) 中的第一个包含关系由通过 ε- 法向量的极限给出的基本法向量的定义直接可得. 为证 (ii) 中的第二个包含关系, 注意到集合 $\mathcal{E}(f,\Omega,\Theta)$ 可表示为逆像

$$\mathcal{E}(f,\Omega,\Theta)=g^{-1}(\Theta),\quad 其中\ \ g(x,z):=f_\Omega(x)-z;$$

因此可对 $F=g$ 情形应用关于 Asplund 空间中的逆像的基本法向量的定理 3.8. 易见, 在所给的连续性和闭性的假设下, $g(\cdot)\cap\Theta$ 在 $(\bar{x},\bar{z})$ 是内半紧的, 且 $g(\bar{x},\bar{z})\cap\Theta=f(\bar{x})-\bar{z}$. 下面证明, 对上面特殊结构的映射 g 来说, 有

$$\ker\widetilde{D}^*_M g(\bar{x},\bar{z})=\{0\},\quad 且\ \ g^{-1}\ 在\ (f(\bar{x})-\bar{z},\bar{x},\bar{z})\ 是\ \text{PSNC}\ 的.$$

首先验证该条件. 任取 $z^*\in Z^*$ 满足 $0\in\widetilde{D}^*_M g(\bar{x},\bar{z})(z^*)$, 找到 $(x_k,z_k)\to(\bar{x},\bar{z})$, $(u_k^*,v_k^*)\in\widehat{D}^*g(x_k,z_k)(z_k^*)$, 使得

$$x_k\in\Omega,\quad \|(u_k^*,v_k^*)\|\to 0,\quad z_k^*\xrightarrow{w^*}z^*\quad(k\to\infty).$$

由于 $g(x,z)=f_\Omega(x)-z$, 根据 1.62(i) 有

$$\widehat{D}^*g(x_k,z_k)(z_k^*)=\Big(\widehat{D}^*f_\Omega(x_k)(z_k^*),0\Big)+(0,-z_k^*).$$

因此

$$u_k^*\in\widehat{D}^*f_\Omega(x_k)(z_k^*),\quad v_k^*=-z_k^*,\quad k\in\mathbb{N},$$

从而 $\|z_k^*\|\to 0=z^*$, 即 $\ker\widetilde{D}^*_M g(\bar{x},\bar{z})=\{0\}$. 为验证 g^{-1} 在 $(f(\bar{x})-\bar{z},\bar{x},\bar{z})$ 的 PSNC 性质, 用类似的方法, 取 $x_k\in\Omega$,

$$(u_k^*,v_k^*,z_k^*)\in\widehat{N}((x_k,z_k,f(x_k)-z_k);\text{gph}\,g)\ 满足\ (x_k,z_k)\to(\bar{x},\bar{z}),$$

$$\|(u_k^*, v_k^*)\| \to 0, \quad z_k^* \xrightarrow{w^*} 0 \quad (k \to \infty).$$

于是, 根据上面的论证就有 $\|z_k^*\| \to 0$, 这就证明了 g^{-1} 在 $(f(\bar{x}) - \bar{z}, \bar{x}, \bar{z})$ 的 PSNC 性质. 即验证了定理 3.8 中的所有假设, 从而由下式得 (5.38)

$$(u^*, v^*) \in D_N^* g(\bar{x}, \bar{z})(z^*) \Leftrightarrow u^* \in D_N^* f_\Omega(\bar{x})(z^*), v^* = -z^*,$$

而该式由对 $g(x, z) = f_\Omega(x) - z$ 应用定理 1.62(ii) 中的和式法则可得.

现在在其中额外的假设下证 (iii) 中的包含关系 (5.39). 考虑到断言 (i), 有

$$D_N^* f_\Omega(\bar{x})(z^*) = D_M^* f_\Omega(\bar{x})(z^*) = \partial\langle z^*, f_\Omega\rangle(\bar{x}), \quad z^* \in Z^*, \tag{5.40}$$

故 (5.39) 由 (5.38) 可得.

下证 (iv). 先证 (5.39) 中相反的包含关系成立. 任取 $z^* \in -N(f(\bar{x}) - \bar{z}; \Theta)$ 和 $x^* \in \partial\langle z^*, f_\Omega\rangle(\bar{x})$, 欲证 $(x^*, z^*) \in N((\bar{x}, \bar{z}); \mathcal{E}(f, \Omega, \Theta))$. 由基本法向量和次梯度的定义, 有序列 $\varepsilon_{1k} \downarrow 0, \varepsilon_{2k} \downarrow 0, x_k \xrightarrow{\Omega} \bar{x}, z_k \to \bar{z}, x_k^* \xrightarrow{w^*} x^*$和$z_k^* \xrightarrow{w^*} z^*$ $(k \to \infty)$ 使得

$$x_k^* \in \widehat{\partial}_{\varepsilon_{1k}}\langle z^*, f_\Omega\rangle(x_k), \quad -z_k^* \in \widehat{N}_{\varepsilon_{2k}}(f(x_k) - z_k; \Theta), \quad \forall k \in \mathbb{N}.$$

由 ε- 法向量和 ε- 次梯度的定义, 并利用 f_Ω 在 $\bar{x}$ 相对于 Ω 的 w^*- 严格 Lipschitz 连续性 (请比照定理 3.28 的证明), 上面的包含关系蕴涵着

$$(x_k^*, z_k^*) \in \widehat{N}_{\varepsilon_k}((x_k, z_k); \mathcal{E}(f, \Omega, \Theta)), \quad \text{对大的 } k \in \mathbb{N} \text{ 成立},$$

其中 $\varepsilon_k \downarrow 0$ 为某正数列. 对 $k \to \infty$ 取极限 (5.34) 中的反向包含关系. (5.38) 中的反向包含关系在上导数正规性假设下根据基本上导数表示 (5.40) 由 (5.39) 中的方向包含关系可得. △

现在回到本小节考虑的主要最优化问题 (5.23), 并且定义集合

$$\mathcal{E}(\varphi_0, \cdots, \varphi_{m+r}, \Omega) := \Big\{(x, \alpha_0, \cdots, \alpha_{m+r}) \in X \times \mathbb{R}^{m+r+1} \,\Big|\, x \in \Omega, \varphi_i(x) \leqslant \alpha_i,$$
$$i = 0, \cdots, m; \varphi_i(x) = \alpha_i, i = m+1, \cdots, m+r\Big\},$$

它对应于当 $f = (\varphi_0, \cdots, \varphi_{m+r})\colon X \to \mathbb{R}^{m+r+1}, \Theta = \mathbb{R}_-^{m+1} \times \{0\} \subset \mathbb{R}^{m+r+1}$ 时的 (5.37). 基于极点原理, 下面的结果通过广义上图 $\mathcal{E}(\varphi_0, \cdots, \varphi_{m+r}, \Omega)$ 的基本法向量, 在非常广的框架下给出了问题 (5.23) 的必要最优性条件, 并且在关于 $\varphi_i(i = 0, \cdots, m+r)$ Lipschitz 的假设下可用 Lagrange 原理的推广形式来等价地表达. 为方便起见, 下面假设在所考虑的最优解处有 $\varphi_0(\bar{x}) = 0$, 这并不限制一般性.

定理 5.24(推广的 Lagrange 原理) 设 $\bar{x}$ 是问题 (5.23) 的局部最优解, X 是 Asplund 空间. 假设集合 Ω 在 $\bar{x}$ 附近是局部闭的, 对 $i = 0, \cdots, m$, 函数 φ_i 在 $\bar{x}$ 附

近相对于 Ω 是 l.s.c. 的, 且对 $i=m+1,\cdots,m+r$, 函数 φ_i 在该点附近相对于 Ω 是连续的. 则存在 Lagrange 乘子 $(\lambda_0,\cdots,\lambda_{m+r})\in\mathbb{R}^{m+r+1}$, 它们不同时等于 0, 使得

$$(0,-\lambda_0,\cdots,-\lambda_{m+r})\in N((\bar{x},0);\mathcal{E}(\varphi_0,\cdots,\varphi_{m+r},\Omega)) \tag{5.41}$$

成立, 这自动蕴涵着 (5.25) 中的符号和互补松弛条件. 而且, 如果函数 φ_i $(i=1,\cdots,m)$ 在 $\bar{x}$ 附近相对于 Ω 是连续的, 则 (5.41) 还蕴涵着

$$0\in D_N^*\big((\varphi_0,\cdots,\varphi_{m+r})+\Delta(\cdot;\Omega)\big)(\bar{x})(\lambda_0,\cdots,\lambda_{m+r}). \tag{5.42}$$

更进一步, 如果所有函数 φ_i $(i=0,\cdots,m+r)$ 在 $\bar{x}$ 附近相对于集合 Ω 都是 Lipschitz 连续的, 则上导数包含关系 (5.42) 等价于涉及本质 Lagrange 函数的次微分包含关系

$$0\in\partial L_\Omega(\cdot,\lambda_0,\cdots,\lambda_{m+r})(\bar{x}). \tag{5.43}$$

此时必要条件 (5.41) 等价于 (5.25) 和 (5.43) 同时成立.

证明 由于 $\bar{x}$ 是 (5.23) 的局部最优解, 故存在 $\bar{x}$ 的邻域 U, 使得 $\bar{x}$ 是 φ_0 在 U 上相对于 (5.23) 中的约束的最小值点. 简单起见, 设 $\varphi_0(\bar{x})=0$. 考虑 Asplund 空间 $X\times\mathbb{R}^{m+r+1}$ 中的集合

$$\Omega_1:=\mathcal{E}(\varphi_0,\cdots,\varphi_{m+r},\Omega),\quad \Omega_2:=\operatorname{cl}U\times\{0\},$$

并注意到 $(\bar{x},0)$ 是系统 $\{\Omega_1,\Omega_2\}$ 的极点. 事实上, 显然有 $(\bar{x},0)\in\Omega_1\cap\Omega_2$, 且对任何负数列 $\nu_k\uparrow 0$, 根据 (5.23) 中 $\bar{x}$ 的局部最优性, 有 $(\Omega_1-(0,\nu_k,0,\cdots,0))\cap\Omega_2=\varnothing$, $k\in\mathbb{N}$. 考虑到集合 Ω_1 和 Ω_2 在 $(\bar{x},0)$ 都是局部闭的, 且由 $\bar{x}\in\operatorname{int}U$ 和 $0\in\mathbb{R}^{m+r+1}$ 知 Ω_2 在该点是 SNC 的, 利用定理 2.22 中极点原理的确切版本, 得 (5.41) 成立, 且 $(\lambda_0,\cdots,\lambda_{m+r})\neq 0$. 根据 $\Theta=\mathbb{R}_-^{m+1}\times\{0\}$ 情形的引理 5.23(ii), 由 (5.41) 立即得互补松弛和符号条件 (5.25), 并且在 φ_i $(i=m+1,\cdots,m+r)$ 的连续性假设下得到上导数包含关系 (5.42). 如果对所有的 $i=0,\cdots,m+r,\varphi_i$ 在 $\bar{x}$ 附近都是局部 Lipschitz 的, 由于对任何具有有限维像空间的映射来说, 上导数正规性假设都成立, 所以这个定理的等价性陈述由引理 5.23 中的断言 (iii) 和 (iv) 可得. △

进一步, 应用基本法向量、上导数和次梯度的分析法则, 可以得到包含关系 (5.41)~(5.43) 的各种推论. 下面给出利用关于最优化问题 (5.23) 中的费用函数和泛函 (而不是几何) 约束的标准 Lagrange 函数 (5.35) 的次梯度来表达的一些结果.

推论 5.25(Lagrange 函数条件和抽象最大值原理) 设 $\bar{x}$ 是 (5.23) 的局部最优解, X 是 Asplund 空间, 对所有的 $i=0,\cdots,m+r$, 函数 φ_i 在 $\bar{x}$ 附近都是 Lipschitz 连续的, 并假设集合 Ω 在该点附近是局部闭的. 则存在不全为 0 的 Lagrange 乘子

$\lambda_0, \cdots, \lambda_{m+r}$, 使得条件 (5.25) 和

$$0 \in \partial L(\cdot, \lambda_0, \cdots, \lambda_{m+r})(\bar{x}) + N(\bar{x}; \Omega) \tag{5.44}$$

成立. 进一步, 如果集合 Ω 是凸的, 则对某 $x^* \in -\partial L(\cdot, \lambda_0, \cdots, \lambda_{m+r})(\bar{x})$, 有

$$\langle x^*, \bar{x} \rangle = \max \left[\langle x^*, \bar{x} \rangle \middle| x \in \Omega \right]. \tag{5.45}$$

证明 包含关系 (5.44) 由 (5.43) 根据定理 2.33(c) 中的次微分和法则可得. 在凸几何约束的情形下, 根据命题 1.5 中凸集的基本法向量的表示, 这个包含关系蕴涵着最大值条件 (5.45). △

值得注意的是, 推论 5.25 中的第二个断言给出了一个抽象最大值原理, 它直接由通过凸几何约束的法锥来表示的凸结构推得. 还注意到类似推论 5.20 和定理 5.21(iii), 所得的结果蕴涵着由分离约束表示的结果.

转到下一个话题, 注意到在定理 5.21(iii) 得到极小化问题 (5.23) 的下次微分条件时, 利用了基本次梯度集合 $\partial\varphi_i(\bar{x})$ $(i = 0, \cdots, m)$ 和 $\partial\varphi_i(\bar{x}) \cup \partial(-\varphi_i)(\bar{x})$ $(i = m+1, \cdots, m+r)$, 它们是经典严格导数的非光滑推广. 尽管这种类型的结果对无限维空间中的一般等式约束来说似乎是不可改进的, 但是利用非光滑费用和不等式约束函数的通常 —— 不是严格 ——Fréchet 导数的推广, 在某些情况下可得到更精细的 (一般来说是独立的) 结果. 这个方向的一些结果已在定理 5.19 中通过 Fréchet 上次微分给出. 现在建立混合型的下次微分条件, 在等式约束函数上利用严格导数的次梯度推广, 而在目标及不等式约束函数上利用通常 Fréchet 导数的次梯度推广.

为此, 回顾非光滑分析中关于函数和集合的凸方向逼近的一些概念. 给定在 $\bar{x}$ 有限的 $\varphi\colon X \to \overline{\mathbb{R}}$, 增广实值函数

$$d^+\varphi(\bar{x}; h) := \limsup_{z \to h,\, t \downarrow 0} \frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t} \tag{5.46}$$

称为 φ 在 $\bar{x}$ 沿方向 h 的上 Dini-Hadamard 方向导数. 如果 φ 在 $\bar{x}$ 附近是 Lipschitz 的, 在 (5.46) 中可固定 $z = h$. 如果函数 $p(\bar{x}; \cdot)\colon X \to \overline{\mathbb{R}}$ 是凸 l.s.c. 和正齐次的, 且满足对任意 $h \in X$, 有 $p(\bar{x}; h) \geqslant d^+\varphi(\bar{x}; h)$, 则称 p 为 φ 在 $\bar{x}$ 的一个上凸逼近. 此时在凸分析意义下, $p(\bar{x}; \cdot)$ 在 $h = 0$ 处的次微分称为 φ 在 $\bar{x}$ 的 p- 次微分, 记为

$$\partial_p\varphi(\bar{x}) := \partial p(\bar{x}; 0) = \{x^* \in X^* \mid \langle x^*, h \rangle \leqslant p(\bar{x}; h),\ \forall h \in X\}. \tag{5.47}$$

值得注意的是, 次微分 (5.47) 依赖于上凸逼近 $p(\bar{x}; \cdot)$, 所以它不是唯一定义的. 例如, 在 $\mathbb{R}$ 上的函数 $\varphi(x) = -|x|$ 在 $\bar{x} = 0$ 有由满足对任意的 $\gamma \in [-1, 1], p(0; h) = \gamma h$ 的 p 给出的一族上凸逼近. 由 2.5.2A 小节知, Clarke 广义方向导数 $\varphi^\circ(\bar{x}; h)$ 自动给出了任意局部 Lipschitz 函数 φ 的一个上凸逼近. 然而, 这个逼近不一定是最好的,

如上例 $\varphi(x)=-|x|$. 还注意到若 φ 在 $\bar{x}$ 是 Gâteaux 可微的, 则 $p(\bar{x};h)=\langle\nabla\varphi(\bar{x}),h\rangle$ 是 φ 的一个上凸逼近, 即 p- 次微分是函数在一参考点处的通常 (不是严格的) 导数的非光滑推广. 对依照 (5.47) 中格式定义的凸值次微分, 相应于有上凸逼近的不同的特殊函数类, 这种思想有各种有效的实现, 这是初创非光滑分析时的发展路线; 参见 1.4.1 小节中的评注和参考文献.

再回顾在 1.1.2 小节给出的集合 $\Omega\subset X$ 在 $\bar{x}\in\Omega$ 的相依锥, 定义为

$$T(\bar{x};\Omega):=\mathop{\mathrm{Limsup}}_{t\downarrow 0}\frac{\Omega-\bar{x}}{t}. \tag{5.48}$$

这是一个非空闭锥, 对凸集 Ω 它简化为经典的切锥, 而 (5.48) 一般来说是非凸的. 注意到 $\forall\varepsilon\geqslant 0$, 有

$$\widehat{N}_\varepsilon(\bar{x};\Omega)\subset\left\{x^*\in X^*\,\middle|\,\langle x^*,v\rangle\leqslant\varepsilon\|v\|,\forall v\in T(\bar{x};\Omega)\right\}. \tag{5.49}$$

而且, 如果 X 是有限维的, 则 (5.49) 作为等式成立. 因此在这种情形下, 根据 (5.49) 中的等式关系, Fréchet 法锥 $\widehat{N}(\bar{x};\Omega)$ 是相依锥 (5.48) 的极/对偶.

定理 5.26(局部极小值的混合次微分条件) 设 $\bar{x}$ 是问题 (5.23) 的局部最优解, X 是 Asplund 空间, 集合 Ω 在 $\bar{x}$ 附近是局部闭的, 所有的函数 φ_i 在该点附近都是局部 Lipschitz 的. 还假设存在 $T(\bar{x};\Omega)$ 的一个凸闭子锥 M 满足 $M^*\subset N(\bar{x};\Omega)$, 假设函数 $\varphi_i,i\in I(\bar{x})\cup\{0\}$, 在 $\bar{x}$ 有上凸逼近, 它们在 M 中的某点是连续的, 记 $\vartheta(x):=\|(\varphi_{m+1}(x),\cdots,\varphi_{m+r}(x))\|$, 并假设这个函数在 $\bar{x}$ 有一上凸逼近, 它的次微分 (5.47) 包含在 $\partial\vartheta(\bar{x})$ 中. 则对 $i\in I(\bar{x})\cup\{0\}$, 存在 Lagrange 乘子 $\lambda_i\geqslant 0$ 和不同时等于 0 的 $(\lambda_{m+1},\cdots,\lambda_{m+r})\in\mathbb{R}^r$, 使得

$$0\in\sum_{i\in I(\bar{x})\cup\{0\}}\lambda_i\partial_p\varphi_i(\bar{x})+\partial\left(\sum_{i=m+1}^{m+r}\lambda_i\varphi_i\right)(\bar{x})+N(\bar{x};\Omega), \tag{5.50}$$

这里, $\partial_p\varphi_i(\bar{x})$ 表示对应于 φ_i 的上凸逼近 $p_i(\bar{x};\cdot)$ 的次微分 (5.47), $i\in I(\bar{x})\cup\{0\}$.

证明 首先考虑当 $f:=(\varphi_{m+1},\cdots,\varphi_{m+r})\colon X\to\mathbb{R}^r$ 在 $\bar{x}$ 相对于 Ω 是度量正则的情形. 于是定理 5.16 保证了对某 $\mu>0,\bar{x}$ 是其中定义的无约束最小化问题 (UP) 的局部最优解. 利用 (UP) 中费用函数的形式和上凸逼近的定义以及关于 $M\subset T(\bar{x};\Omega)$ 的凸性假设, 由凸规划标准的分离论证及凸分析中标准次微分公式可得, 存在 $\lambda_i\geqslant 0,i\in I(\bar{x})\cup\{0\}$, 且 $\displaystyle\sum_{I(\bar{x})\cup\{0\}}\lambda_i=1$, 使得

$$0\in\sum_{i\in I(\bar{x})\cup\{0\}}\lambda_i\partial_p\varphi_i(\bar{x})+\mu\partial_p\vartheta(\bar{x})+M^*.$$

考虑到 $M^*\subset N(\bar{x};\Omega),\partial_p\vartheta(\bar{x})\subset\partial\vartheta(\bar{x})$, 再应用推论 3.43 中的链式法则于当 $\psi(y):=\|y\|$ 时复合 $\vartheta=(\psi\circ f)$ 的基本次微分, 则在度量正则性的假设下得 (5.50) 成立.

如果 f 在 $\bar{x}$ 相对于 Ω 不是度量正则的, 则 $f_\Omega = f + \Delta(\cdot, \Omega)$ 在定义 1.47(ii) 的意义下在 $\bar{x}$ 附近不是度量正则的. 根据定理 4.18(c) 有 $\ker\widetilde{D}^*_M f_\Omega(\bar{x}) \neq \{0\}$ 或 f_Ω^{-1} 在 $(f(\bar{x}), \bar{x})$ 不是 PSNC 的. 由于 f_Ω 的像空间是有限维的, 则后面的 PSNC 条件自动成立, 故度量正则性的缺失意味着存在非零的 $y^* \in \mathbb{R}^r$, 使得 $0 \in \widetilde{D}^*_M f_\Omega(\bar{x})(y^*)$. 余下的证明由定理 5.17 中的证明可得. △

注意到即使在只有不等式约束的情形下, 定理 5.26 和依据基本次梯度的前面的结果一般来说是独立的. 特别地, 可验证注 5.22 中考虑的最后一个例子中的函数 $\varphi(x_1, x_2) = \big||x_1| + x_2\big| + x_2$ 在 $\bar{x} = 0$ 没有上凸逼近, 它的次微分是基本次微分 $\partial\varphi(0)$ 的真子集. 另一方面, 在一维最优化问题:

$$\begin{cases} \min\ \varphi_0(x) := x \\ \text{s.t.}\ \ \varphi_1(x) := x^2\sin(1/x) \leqslant 0, \quad x \neq 0, \quad \varphi_1(0) = 0 \end{cases}$$

中由定理 5.26 可排除点 $\bar{x} = 0$ 的最优性; 但由于 $\partial\varphi_1(0) = [-1, 1]$, 根据基本次梯度而得的必要条件不起作用.

这一小节最后的结果是关于带有非 Lipschitz 数据的问题 (5.23) 的下次微分必要最优性条件. 对这样问题之前所得的结果是用图像集合和上图集合的法向量来表达, 在缺少 Lipschitz 假设下不能简化为相应函数的次梯度. 下面将得到包含 (5.23) 中费用和约束函数的 Fréchet 次梯度以模糊次微分形式给出的非 Lipschitz 问题的新的必要条件. 为此需要下面的引理, 它是在半 Lipschitz 假设下定理 2.33(b) 给出的 (强) 模糊和法则的弱的非 Lipschitz 情形的相应结果. 注意到这个结果涉及 X^* 中原点的弱 * 邻域, 而不是如定理 2.33(b) 中的小的对偶球. 这个引理根据推论 2.29 中的稠密性结果通过利用下/极小卷积的性质可得; 完整的证明和更多讨论请参见 Fabian[414, 415].

引理 5.27(弱模糊和法则) 设 X 是 Asplund 空间, $\varphi_1, \cdots, \varphi_n$ 是 X 上的增广实值 l.s.c. 函数. 则对任意 $\bar{x} \in X, \varepsilon > 0, x^* \in \widehat{\partial}(\varphi_1 + \cdots + \varphi_n)(\bar{x})$, 及 X^* 中原点的任意弱 * 邻域 V^*, 存在 $x_i \in \bar{x} + \varepsilon\mathbb{B}$ 和 $x_i^* \in \widehat{\partial}\varphi_i(x_i)$, 使得 $|\varphi_i(x_i) - \varphi_i(\bar{x})| \leqslant \varepsilon\ (i = 1, \cdots, n)$ 和

$$x^* \in \sum_{i=1}^n x_i^* + V^*$$

成立.

现在已准备好建立非 Lipschitz 泛函约束的问题 (5.23) 的局部最优解的 Lagrange 乘子法则的一个弱近似版本.

定理 5.28(非 Lipschitz 问题的弱次微分最优性条件) 设 $\bar{x}$ 是 Asplund 空间 X 中问题 (5.23) 的局部最优解. 假设对 $i = 0, \cdots, m$, 函数 φ_i 在 $\bar{x}$ 附近是 l.s.c. 的, 对 $i = m+1, \cdots, m+r$, 函数 φ_i 在该点附近是连续的, 并假设集合 Ω 在 $\bar{x}$ 附近是

局部闭的. 则 $\forall \varepsilon > 0$ 和 X^* 中原点的任何弱 * 邻域 V^*, 都存在

$$x_i \in \bar{x} + \varepsilon\mathbb{B} \text{ 满足 } \left|\varphi_i(x_i) - \varphi_i(\bar{x})\right| \leqslant \varepsilon \quad (i = 0, \cdots, m+r),$$

$$x_i^* \in \widehat{\partial}\varphi_i(x_i) \quad (i = 0, \cdots, m),$$

$$x_i^* \in \widehat{\partial}\varphi_i(x_i) \cup \widehat{\partial}(-\varphi_i)(x_i) \quad (i = m+1, \cdots, m+r),$$

$$\widehat{x}^* \in \widehat{N}(\widehat{x}; \Omega), \text{ 其中 } \widehat{x} \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B}), \text{ 以及}$$

$$\lambda_i \geqslant 0 \ (i = 0, \cdots, m+r) \text{ 且 } \sum_{i=0}^{m+r} \lambda_i = 1$$

满足关系

$$0 \in \sum_{i=0}^{m+r} \lambda_i x_i^* + \widehat{x}^* + V^*.$$

证明 考虑约束集合

$$\Omega(x) := \begin{cases} \{x \in X | \ \varphi_i(x) \leqslant 0\} & (i = 1, \cdots, m), \\ \{x \in X | \ \varphi_i(x) = 0\} & (i = m+1, \cdots, m+r) \end{cases}$$

并注意到原始约束问题 (5.23) 显然等价于具有“无限惩罚”的无约束问题:

$$\min \varphi_0(x) + \delta(x; \Omega_1 \cap \cdots \cap \Omega_{m+r} \cap \Omega), \quad x \in X.$$

根据广义 Fermat 原理和上面优化问题中费用函数的结构有

$$0 \in \widehat{\partial}\left(\varphi_0 + \sum_{i=1}^{m+r} \delta(\cdot; \Omega_i) + \delta(\cdot; \Omega)\right)(\bar{x}).$$

任取 $\varepsilon > 0$ 和 X^* 中原点的一弱 * 邻域 V^*, 再对上面的和式应用引理 5.27, 可找到

$$x_0^* \in \widehat{\partial}\varphi_0(x_0) \text{ 满足 } \|(x_0, \varphi_0(x_0)) - (\bar{x}, \varphi_0(\bar{x}))\| \leqslant \varepsilon,$$

$$\widetilde{x}^* \in \widehat{N}(\widehat{x}; \Omega) \text{ 满足 } \widehat{x} \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B}),$$

$$\widetilde{x}_i^* \in \widehat{N}(\widetilde{x}_i; \Omega_i) \text{ 满足 } \widetilde{x}_i \in \Omega_i \cap (\bar{x} + (\varepsilon/2)\mathbb{B}) \ (i = 1, \cdots, m+r)$$

使下式成立

$$0 \in x_0^* + \sum_{i=1}^{m+r} \widetilde{x}_i^* + \widetilde{x}^* + \left(\frac{1}{m+r+1}\right)V^*.$$

考虑到集合 Ω_i 的结构, 分别讨论下面两种情形.

情形 (a). 假设此时或者存在 $i \in \{1, \cdots, m\}$ 和 $\lambda \neq 0$ 满足 $(0, \lambda) \in N((\widetilde{x}_i, 0); \operatorname{epi}\varphi_i)$, 或者存在 $i \in \{m+1, \cdots, m+r\}$ 满足 $0 \in \partial\varphi_i(\widetilde{x}_i) \cup \partial(-\varphi_i)(\widetilde{x}_i)$. 若对某个 $i \in \{1, \cdots, m\}$ 上述相应结论成立, 则根据 Asplund 空间中定理 2.35 中的基本法锥表示, 可找到 $(x_i, \alpha_i) \in \operatorname{epi}\varphi_i, (x_i^*, -\lambda_i) \in \widehat{N}((x_i, \alpha_i); \operatorname{epi}\varphi_i)$, 使得

$$\|x_i - \widetilde{x}_i\| \leqslant \frac{\varepsilon}{2}, \quad \lambda_i > 0, \quad 且 \quad x_i^* \in \lambda_i V^*.$$

易见, 由 $\alpha_i \geqslant \varphi_i(x_i)$ 得 $\widehat{N}((x_i, \alpha_i); \operatorname{epi}\varphi_i) \subset \widehat{N}((x_i, \varphi_i(x_i)); \operatorname{epi}\varphi_i)$, 因此 $(x_i^*, -\lambda_i) \in \widehat{N}((x_i, \varphi_i(x_i)); \operatorname{epi}\varphi_i)$. 从而在这种情形下有包含关系

$$x_i^*/\lambda_i \in \widehat{\partial}\varphi_i(x_i), \quad x_i^*/\lambda_i \in V^*.$$

这蕴涵着对参数指标 i, 在 $\widehat{x}^* = 0, \lambda_i = 1$(其他 λ_i 都是 0) 时, 这个定理中所要求的关系都成立.

继续 (a) 的论证, 现在考虑指标 $i \in \{m+1, \cdots, m+r\}$ 的情形. 利用定理 2.34(b) 关于在 Asplund 空间上连续函数的基本次微分表示, 找到 $(x_i, x_i^*) \in X \times X^*$, 使得

$$x_i \in \widetilde{x}_i + (\varepsilon/2)\mathbb{B}, \quad x_i^* \in \left[\widehat{\partial}\varphi_i(x_i) \cup \widehat{\partial}(-\varphi_i)(x_i)\right] \cap V^*.$$

这也蕴涵着对参考指标 i, $\widehat{x}^* = 0$, $\lambda_i = 1$, 而其他所有 λ_i 全等于 0 时, 这个定理的结论成立.

情形 (b). 设 (a) 中的假设不成立. 首先考虑对应着等式约束的一个指标 $i \in \{m+1, \cdots, m+r\}$, 即 $\Omega_i = \{x \in X | \varphi_i(x) = 0\}$ 的情形. 对

$$\Lambda_1 := \operatorname{gph}\varphi_i, \quad \Lambda_2 := \{(x, \alpha) \in X \times \mathbb{R} \,|\, \alpha = 0\},$$

注意到 $\Omega_i \times \{0\} = \Lambda_1 \cap \Lambda_2$, 其中集合 Λ_2 在 $(\widetilde{x}_i, 0)$ 是 SNC 的, 且推论 3.5 中的规范条件简化为 $0 \notin \partial\varphi_i(\widetilde{x}_i) \cup \partial(-\varphi_i)(\widetilde{x}_i)$. 现应用这个引理中的交公式, 再应用定理 1.80 和 2.40(ii), 它们给出了连续函数上导数的次微分表示, 则得包含关系

$$N(\widetilde{x}_i; \Omega_i) \subset \partial^\infty\varphi_i(\widetilde{x}_i) \cup \partial^\infty(-\varphi_i)(\widetilde{x}_i) \cup \mathbb{R}_+\partial\varphi_i(\widetilde{x}_i) \cup \mathbb{R}_+\partial(-\varphi_i)(\widetilde{x}_i),$$

这里 $\mathbb{R}_+ S := \{\nu s | \nu \geqslant 0, s \in S\}$. 利用定理 2.34(b), 2.35(b) 和 2.38 中法向量和次微分的表示, 这个包含关系蕴涵着对任意的 $i = m+1, \cdots, m+r$, 存在 $x_i \in \widetilde{x}_i + (\varepsilon/2)\mathbb{B}, \nu_i \geqslant 0$ 和

$$x_i^* \in \widehat{\partial}\varphi_i(x_i) \cup \widehat{\partial}(-\varphi_i)(x_i) \quad 满足 \quad \nu_i x_i^* \in \widetilde{x}_i^* + \frac{1}{m+r+1}V^*.$$

接下来考虑对应着不等式约束的一个指标 $i \in \{1, \cdots, m\}$, 即 $\Omega_i := \{x \in X | \varphi_i(x) \leqslant 0\}$ 的情形. 令

$$\Lambda_1 := \operatorname{epi}\varphi_i, \quad \Lambda_2 := \{(x, \alpha) \in X \times \mathbb{R} \,|\, \alpha = 0\},$$

则 $\Omega_i \times \{0\} = \Lambda_1 \cap \Lambda_2$. 注意到推论 3.5 的假设在 $(\widetilde{x}_i, 0)$ 成立, 这是由于 Λ_2 在该点是 SNC 的, 并且这个推论的规范条件简化为

$$(0, \lambda) \in N\big((\widetilde{x}_i, 0); \operatorname{epi} \varphi_i\big) \Rightarrow \lambda = 0.$$

因此, 取上面的 $(\widetilde{x}_i, \widetilde{x}_i^*)$ 满足

$$\widetilde{x}_i^* \in \widehat{N}(\widetilde{x}_i; \Omega_i) \subset N(\widetilde{x}_i; \Omega_i), \quad \widetilde{x}_i \in \Omega_i \cap \big(\bar{x} + (\varepsilon/2)\mathbb{B}\big),$$

则可找到 $\widetilde{\nu}_i \geqslant 0$ 使得 $(\widetilde{x}_i, -\widetilde{\nu}_i) \in N((\widetilde{x}_i, 0); \operatorname{epi} \varphi_i)$. 于是利用定理 2.35(b) 中基本法向量的极限表示, 当 $(\widehat{x}_i, \alpha_i)$ 充分接近 $(\widetilde{x}_i, 0)$ 时可用元素 $(\widehat{x}_i^*, -\widehat{\nu}_i) \in \widehat{N}\big((\widehat{x}_i, \alpha_i); \operatorname{epi} \varphi_i\big)$ 在 $X^* \times \mathbb{R}$ 的弱 * 拓扑下逼近 $(\widetilde{x}_i^*, \widetilde{\nu}_i)$. 不失一般性, 可假设 $\alpha_i = \varphi_i(\widehat{x}_i)$; 请比照 (a). 如果 $\widehat{\nu}_i \neq 0$, 置 $x_i := \widehat{x}_i, \nu_i := \widehat{\nu}_i$ 和 $x_i^* := \widehat{x}_i^*/\nu_i$ 可以得到

$$x_i^* \in \widehat{\partial}\varphi_i(x_i) \text{ 满足 } \nu_i x_i^* \in \widetilde{x}_i^* + \frac{1}{m+r+1} V^*.$$

如果 $\widehat{\nu}_i = 0$, 利用引理 2.37 找到 $(\widehat{x}_i^*, \widehat{x}_i)$ 在 $X^* \times X$ 的范数拓扑下的一个强逼近 $(\nu_i x_i^*, x_i)$, 使得 $\nu_i \geqslant 0$ 且 $x_i^* \in \widehat{\partial}\varphi_i(x_i)$.

结合上面的关系, 有

$$\|\widehat{x} - \bar{x}\| \leqslant \varepsilon, \quad \|x_i - \bar{x}\| \leqslant \varepsilon \ \ (i = 0, \cdots, m+r), \quad |\varphi_0(x_0) - \varphi_0(\bar{x})| \leqslant \varepsilon,$$

$$\widetilde{x}^* \in \widehat{N}(\widehat{x}; \Omega), \quad x_i^* \in \widehat{\partial}\varphi_i(x_i) \quad (i = 0, \cdots, m),$$

$$x_i^* \in \widehat{N}(\widehat{x}; \Omega), \quad x_i^* \in \widehat{\partial}\varphi_i(x_i) \quad (i = 0, \cdots, m),$$

$$x_i^* \in \widehat{\partial}\varphi_i(x_i) \cup \widehat{\partial}(-\varphi_i)(x_i) \quad (i = m+1, \cdots, m+r),$$

满足$0 \in x_0^* + \sum_{i=1}^{m+r} \nu_i x_i^* + \widehat{x}^* + V^*$, 其中对所有的 $i = 1, \cdots, m+r$ 有 $\nu_i \geqslant 0$. 现令 $\lambda := 1 \Big/ \left(1 + \sum_{i=1}^{m+r} \nu_i\right)$, $\widehat{x}^* := \lambda \widetilde{x}^*, \lambda_0 := \lambda$ 和对 $i = 1, \cdots, m+r$, $\lambda_i := \lambda \nu_i$, 除了

$$|\varphi_i(x_i) - \varphi_i(\bar{x})| \leqslant \varepsilon, \quad i = 1, \cdots, m, \tag{5.51}$$

得到了这个定理的其他结论. 注意到对 $i = m+1, \cdots, m+r$, 估计 (5.51) 由 φ_i 的连续性自动得出. 对 $i = 1, \cdots, m$, 当 φ_i 被假设为在 $\bar{x}$ 附近仅仅是 l.s.c. 的, 则情况就不同了. 注意到, 如果 $\varphi_i(\bar{x}) = 0$, 则对这个 i, (5.51) 由 φ_i 的下半连续性直接可得. 否则若对某个 $i \in \{1, \cdots, m\}, \varphi_i(\bar{x}) < 0$, 则用 $\phi_i(x) := \varphi_i(x) - \varphi_i(\bar{x}) \leqslant 0$ 来取代这个约束, 并注意到 $\bar{x}$ 是新问题的最优解并满足 $\phi_i(\bar{x}) = 0$ 且 $\widehat{\partial}\phi_i(\bar{x}) = \widehat{\partial}\varphi_i(\bar{x})$. 这就完全地证明了 (5.51) 成立, 从而完成了定理的证明. △

5.1.4 约束问题的次优性条件

这一小节研究无限维空间中数学规划问题的次优性条件, 这意味着不假设最优解的存在性, 所得条件针对的是次优 (ε- 最优) 解, 它们总是存在的. 次优解对无限维最优化问题来说特别重要, 因为在该情况下, 最优解的存在性的要求相当苛刻. 正如 L. C. Young 指出的那样, 必要最优性条件的任何理论在最优解的存在性被澄清前都是“不成熟/肤浅”的. 作为最基本的初始动力, 这促进变分法和最优控制问题中的广义曲线/松弛控制理论的发展, 以便能自动保证最优解的存在; 更多细节和讨论参见第 6 章. 然而, 在刚提到的无限维最优化领域中建立的这些方法本质上基于由可微性和相关方程控制的连续时间动态约束的特殊特征. 这不适用于无限维空间中的一般最优化问题. 在一般最优化问题中, 避免最优解的存在性困难的自然方法是阐明“几乎”最优 (即次最优) 解“几乎”满足最优性必要条件. 从实践的观点来看, 这与必要最优性条件有大致相同的效果和应用.

下面在初始数据的 Lipschitz 和非 Lipschitz 假设下, 建立具有等式和不等式约束的非可微规划问题 (5.23) 的次微分型必要最优性条件. 类似结果还可对 (5.12) 类型具有算子约束的更一般问题得到, 简洁起见, 本小节没有考虑.

首先研究具有非 Lipschitz 数据的问题 (5.23) 的次最优性条件. 下面的结果类似定理 5.28, 唯一本质的区别在于所得的弱次优性条件不包括对由 l.s.c. 函数给出的不等式约束的结论 (5.51). 其证明也类似定理 5.28 的证明, 但稍微复杂一些, 其中对相应的无约束问题应用的是定理 2.28 中下次微分变分原理, 而不是 Fermat 驻点原理.

回顾一下, 最优化问题 (5.23) 的可行解是指满足所有约束的那些 x, $\inf\varphi_0$ 是指费用函数相对于 (5.23) 所有可行解的下确界. 这里总是假设 $\inf\varphi_0 > -\infty$. 自然地, 称一个点 x 是 (5.23) 的一个 ε- 最优解, 如果它对这个问题是可行的且满足

$$\varphi_0(x) \leqslant \inf\varphi_0 + \varepsilon.$$

定理 5.29(非 Lipschitz 问题的弱次优性条件) 设 X 是 Asplund 空间, V^* 是 X^* 中原点的任一弱 * 邻域. 假设 Ω 是闭的, $\varphi_0,\cdots,\varphi_m$ 是 l.s.c. 的, 假设对所有充分小的 $\varepsilon>0, \varphi_{m+1},\cdots,\varphi_{m+r}$ 在 (5.23) 的 ε- 最优解集上是连续的. 则存在 $\bar\varepsilon>0$, 使得对任意的 $0<\varepsilon<\bar\varepsilon$ 和 (5.23) 的任意 ε^2- 最优解 $\bar x$, 存在 (x_i, x_i^*, λ_i) 满足条件:

$$x_i \in \bar x + \varepsilon\mathbb{B} \quad (i=0,\cdots,m+r) \quad \text{且} \quad |\varphi_0(x_0)-\varphi_0(\bar x)| \leqslant \varepsilon,$$

$$x_i^* \in \widehat\partial\varphi_i(x_i) \quad (i=0,\cdots,m),$$

$$x_i^* \in \widehat\partial\varphi_i(x_i) \cup \widehat\partial(-\varphi_i)(x_i) \quad (i=m+1,\cdots,m+r),$$

$$\widehat x^* \in \widehat N(\widehat x;\Omega), \quad \text{其中} \ \widehat x \in \Omega\cap(\bar x+\varepsilon\mathbb{B}),$$

$$\lambda_i \geqslant 0 \quad (i=0,\cdots,m+r) \text{ 满足 } \sum_{i=0}^{m+r}\lambda_i = 1, \text{ 且}$$

$$0 \in \sum_{i=0}^{m+r}\lambda_i x_i^* + \widehat{x}^* + V^*.$$

证明 对任意的 $v \in X$ 和 $\gamma > 0$, 考虑由

$$V^*(v;\gamma) := \{x^* \in X^* \mid |\langle x^*, v\rangle| < \gamma\}$$

给出的 X^* 中原点的一族弱 * 邻域, 它们组成了一个弱 * 拓扑基. 任取一个弱 * 邻域 V^*, 找到 $\bar{\gamma} > 0, p \in \mathbb{N}$ 和 $v_j \in X$ 满足 $\|v_j\| = 1, 1 \leqslant j \leqslant p$, 使得

$$\bigcap_{j=1}^{p} V^*(v_j; 2\bar{\gamma}) \subset V^*.$$

下证对任意满足 $0 < \varepsilon < \bar{\varepsilon} := \min\{\bar{\gamma}, 1\}$ 的 ε, 该定理中的结论成立. 事实上, 任取可行点 $\bar{x}$ 满足 $\varphi_0(\bar{x}) < \inf\varphi_0 + \varepsilon^2$, 并且找到 $\eta \in (0,\varepsilon)$, 使得 $\varphi_0(\bar{x}) < \inf\varphi_0 + (\varepsilon-\eta)^2$ 成立. 考虑约束集合 Ω_i, 它们与在定理 5.28 证明中的定义相同, 注意到对函数

$$\varphi(x) := \varphi_0(x) + \delta(x;\Omega_1) + \cdots + \delta(x;\Omega_{m+r}) + \delta(x;\Omega), \quad x \in X,$$

有 $\varphi(x) < \inf_X \varphi + (\varepsilon-\eta)^2$. 于是利用定理 2.28(b) 中的下次微分变分原理, 找到 (5.23) 的一可行解 $u \in X$ 和 $u^* \in \widehat{\partial}\varphi(u)$ 满足 $\|u-\bar{x}\| < \varepsilon - \eta$ 和

$$\|u^*\| < \varepsilon - \eta < \bar{\gamma}, \quad \varphi_0(u) < \inf\varphi_0 + (\varepsilon-\eta)^2 < \inf\varphi_0 + \varepsilon - \eta,$$

这蕴涵着 $|\varphi_0(u) - \varphi_0(\bar{x})| < \varepsilon - \eta$.

现取 $\gamma := \bar{\gamma}/(m+r+1)$, 并考虑 $0 \in X^*$ 的弱 * 邻域

$$\widehat{V}^* := \bigcap_{j=1}^{p} V^*(v_j;\gamma).$$

对邻域 $\widehat{V}^*$ 和常数 η, 应用引理 5.27 中的弱模糊和法则于 $u^* \in \widehat{\partial}\varphi(u)$, 然后仿效定理 5.28 的证明, 即可得此定理的所有结论. △

下一结果在恰当的约束规范下给出了具有部分 Lipschitz 数据的问题的规范/标准形式的强次优性条件, 下面对问题 (5.23) 的任意可行解 x, 使用记号

$$I(x) := \{i \in \{1,\cdots,m+r\} | \varphi_i(x) = 0\}, \quad \Lambda(x) := \{\lambda_i \geqslant 0 | i \in I(x)\}.$$

定理 5.30(约束规范下的强次优性条件) 设 X 是 Asplund 空间, $\varepsilon > 0$. 假设 φ_0 是 l.s.c. 的, Ω 是闭的, 在 (5.23) 的 ε- 最优解集上或者 φ_0 是 SNEC 的, 或者

Ω 是 SNC 的. 还假设在此集合上函数 $\varphi_1, \cdots, \varphi_{m+r}$ 在 x 附近是局部 Lipschitz 的, 而且下面的规范条件成立:

$$\text{如果}\ \ x_\infty^* \in \partial^\infty \varphi_0(x), \quad x^* \in N(x;\Omega), \quad x_i^* \in \partial\varphi_i(x), \quad i \in \{1,\cdots,m\} \cap I(x),$$

$$x_i^* \in \partial\varphi_i(x) \cup \partial(-\varphi_i)(x), \quad i \in \{m+1,\cdots,m+r\}, \quad \lambda_i \in \Lambda(x), \ \text{且}$$

$$x_\infty^* + \sum_{i \in I(x)} \lambda_i x_i^* + x^* = 0,$$

则 $x_\infty^* = x^* = 0$ 且 $\lambda_i = 0$ 对任意的 $i \in I(x)$ 成立.

在这些假设下有如下的次优性条件成立: 对 (5.23) 的任意 ε- 最优解 $\bar{x}$ 和任意的 $\nu > 0$, 存在这个问题的一个 ε- 最优解 $\widehat{x}$ 使得 $\|\widehat{x} - \bar{x}\| \leqslant \nu$, 并且估计

$$\left\| \widehat{x}_0^* + \sum_{i \in I(\widehat{x})} \widehat{\lambda}_i \widehat{x}_i^* + \widehat{x}^* \right\| \leqslant \frac{\varepsilon}{\nu}$$

对某 $\widehat{x}_0^* \in \partial\varphi_0(\widehat{x}), \widehat{x}^* \in N(\widehat{x};\Omega), \widehat{\lambda}_i \in \Lambda(\widehat{x})$,

$$\widehat{x}_i^* \in \partial\varphi_i(\widehat{x}), \quad i \in \{1,\cdots,m\} \cap I(\widehat{x}),$$

$$\widehat{x}_i^* \in \partial\varphi_i(\widehat{x}) \cup \partial(-\varphi_i)(\widehat{x}), \quad i = m+1,\cdots,m+r$$

成立. 反之, 如果对 Banach 空间 X 中的任意在其有效域上凹的 l.s.c. 函数 $\varphi_0 \colon X \to \mathbb{R}$, 上面的次优条件对极小化 φ_0 的问题都成立, 则 X 必为 Asplund 的.

证明 如定理 5.29 的证明, 考虑惩罚过的函数

$$\varphi(x) := \varphi_0(x) + \delta(x;\Omega_1) + \cdots + \delta(x;\Omega_{m+r}) + \delta(x;\Omega), \quad x \in X.$$

注意到 $\bar{x}$ 是极小化 φ 的无约束问题的一个 ε- 最优解. 对给定的 $\nu > 0$, 应用下次微分变分原理, 找到原始问题 (5.23) 的一个 ε- 最优解和 $\widehat{x}^* \in \partial\varphi(\widehat{x})$ 满足估计 $\|\widehat{x} - \bar{x}\| \leqslant \nu$ 和 $\|\widehat{x}^*\| \leqslant \varepsilon/\nu$. 然后由次微分等式

$$\partial\varphi(\widehat{x}) = \partial\left(\varphi_0 + \sum_{i \in I(\widehat{x})} \delta(\cdot;\Omega_i) + \delta(\cdot,\Omega)\right)(\bar{x}),$$

利用在推论 3.85 中所得约束集合 Ω_i 的 SNC 性质的有效次微分条件, 在所给的假设下对上面的和应用定理 3.36 中的基本次微分和法则. 当所有的 φ_i 都是 Lipschitz 连续的, 考虑定理 5.28(b) 的证明中集合 Ω_i 的基本法向量的表示, 即得想要的次优性条件.

余下来证此定理的反向陈述. 设 X 是 Banach 空间, $\varphi\colon X\to\mathbb{R}$ 是任一凹连续函数, 根据 φ 的连续性, 对任意的 $\bar{x}\in X,\varepsilon>0$, 存在 $0<\varepsilon_1<\varepsilon$, 使得 $\varphi(\bar{x})<\varphi(x)+2\varepsilon$ 对任意的 $x\in\bar{x}+\varepsilon_1\mathbb{B}$ 成立. 考虑无约束最优化问题:

$$\min\ \varphi_0(x):=\varphi(x)+\delta(x;\bar{x}+\varepsilon_1\mathbb{B}),\quad x\in X.$$

对这个问题应用此定理的次优性条件, 找到 $\widehat{x}\in\bar{x}+(\varepsilon/2)\mathbb{B}$ 满足 $\partial\varphi_0(\widehat{x})=\partial\varphi(\widehat{x})\neq\varnothing$. 根据 Banach 空间上任意连续函数的基本次微分表示 (见定理 1.89)

$$\partial\varphi(\widehat{x})=\operatorname*{Limsup}_{x\to\widehat{x},\ \epsilon\downarrow 0}\widehat{\partial}_\epsilon\varphi(x),$$

对任意 $\epsilon>0$, 存在 $x_\epsilon\in\bar{x}+\varepsilon\mathbb{B}$ 满足 $\widehat{\partial}_\epsilon\varphi(x_\epsilon)\neq\varnothing$. 这蕴涵着, 对任意凹连续函数 $\varphi\colon X\to\mathbb{R}$ 和任意 $\epsilon>0$, 点集 $\{x\in X|\widehat{\partial}_\epsilon\varphi(x)\neq\varnothing\}$ 在 X 中是稠密的, 因此根据推论 2.29(也可参见它之后的讨论), 空间 X 必为 Asplund 空间. △

如果 φ_0 在 (5.23) 的 ε- 最优解集上是 Lipschitz 连续的, 则 φ_0 自然是 SNC 的, 且 $\partial^\infty\varphi_0(x)=\{0\}$. 此时, 定理 5.30 中的规范条件是一个约束规范. 而且, 当函数 $\varphi_1,\cdots,\varphi_{m+r}$ 在这样的 x 点处是严格可微的, 且 $\Omega=X$, 则它简化为经典的 Mangasarian-Fromovitz 约束规范. 由此得到推论.

推论 5.31(Mangasarian-Fromovitz 约束规范下的次优性) 设 X 是 Asplund 空间, 对某 $\varepsilon>0$, φ_0 在 (5.23) 的 ε- 最优解集上是局部 Lipschitz 的, 且 $\Omega=X$. 假设 $\varphi_1,\cdots,\varphi_{m+r}$ 在这样的集合上是严格可微的, 同时满足 M-F 约束规范. 则对 (5.23) 的任意 ε- 最优解 $\bar{x}$ 和任意的 $\nu>0$, 存在这个问题的一个 ε- 最优解 $\widehat{x}$, 一个次梯度 $x_0^*\in\partial\varphi_0(\widehat{x})$ 和乘子 $(\lambda_1,\cdots,\lambda_{m+r})\in\mathbb{R}^{m+r}$, 满足 $\|\widehat{x}-\bar{x}\|\leqslant\nu,\lambda_i\geqslant 0$ 和 $\lambda_i\varphi_i(\widehat{x})=0,i=1,\cdots,m$, 以及

$$\left\|x_0^*+\sum_{i=1}^{m+r}\lambda_i\nabla\varphi_i(\widehat{x})\right\|\leqslant\frac{\varepsilon}{\nu}.$$

证明 因为对严格可微函数来说 $\partial\varphi(x)=\{\nabla\varphi(\bar{x})\}$, 由定理 5.30 直接可得结论成立. △

下面的推论给出了无约束规范的 Lipschitz 问题 (5.23) 的强次优性条件.

推论 5.32(无约束规范的强次优性条件) 设 X 是 Asplund 空间, $\varepsilon>0$. 假设 Ω 是闭的, 所有的 $\varphi_0,\cdots,\varphi_{m+r}$ 在 (5.23) 的 ε- 最优解集上是局部 Lipschitz 的, 则对任意的 $\nu>0$ 和 (5.23) 的任意 ε- 最优解 $\bar{x}$, 存在这个问题的一个 ε- 最优解 $\widehat{x}$, 使得 $\|\widehat{x}-\bar{x}\|\leqslant\nu$ 和

$$\left\|\sum_{i\in I(\widehat{x})\cup\{0\}}\lambda_i\widehat{x}_i^*+\widehat{x}^*\right\|\leqslant\frac{\varepsilon}{\nu},\qquad\sum_{i\in I(\widehat{x})\cup\{0\}}\lambda_i=1$$

对某

$$\lambda_i \geqslant 0 \quad (i \in I(\widehat{x}) \cup \{0\}), \quad \widehat{x}^* \in N(\widehat{x};\Omega), \quad \widehat{x}_0^* \in \partial\varphi_0(\widehat{x}),$$

$$\widehat{x}_i^* \in \partial\varphi_i(\widehat{x}) \quad (i \in \{1,\cdots,m\} \cap I(\widehat{x})),$$

$$\widehat{x}_i^* \in \partial\varphi_i(\widehat{x}) \cup \partial(-\varphi_i)(\widehat{x}) \quad (i = m+1,\cdots,m+r)$$

成立.

证明 首先假设定理 5.30 中的规范条件满足. 于是对某个 $\widehat{\lambda}_i \in \Lambda(\widehat{x})$, 定理的次最优性条件成立. 现在令

$$\lambda := 1 + \sum_{i \in I(\bar{x})} \widehat{\lambda}_i, \quad \lambda_0 := \frac{1}{\lambda} \quad 和 \quad \lambda_i := \frac{\widehat{\lambda}_i}{\lambda}, \quad \forall i \in I(\widehat{x}),$$

则得到具有乘子 $(\lambda_0,\cdots,\lambda_{m+r})$ 的该推论中的所有关系.

然后假设定理 5.30 中的规范条件不满足. 根据 φ_0 的 Lipschitz 连续性知 $x_\infty^* = 0$. 那么可找到问题 (5.23) 的一个 ε- 最优解 $\widehat{x}$, 不全为零的乘子 $\widehat{\lambda}_i \geqslant 0$ $(i \in I(\widehat{x}))$, 以及对偶元素 $x^* \in N(\widehat{x};\Omega)$, $\widehat{x}_i^* \in \partial\varphi_i(\widehat{x})$ $(i \in \{1,\cdots,m\} \cap I(\widehat{x}))$, 与 $\widehat{x}_i^* \in \partial\varphi_i(\widehat{x}) \cup \partial(-\varphi_i)(\widehat{x})$ $(i \in \{m+1,\cdots,m+r\})$ 满足等式

$$\sum_{i \in I(\widehat{x})} \widehat{\lambda}_i \widehat{x}_i^* + x^* = 0.$$

用 $\lambda := \sum\limits_{i \in I(\widehat{x})} \widehat{\lambda}_i > 0$ 除上式, 即得此推论的次最优性条件, 其中 $\lambda_0 := 0$, $\lambda_i := \widehat{\lambda}_i/\lambda \quad (i \in I(\widehat{x}))$, $\widehat{x}^* := x^*/\lambda$, $\widehat{x}_i^*$ 与上述相同. △

注意到如果本节与 5.1.2 小节中的问题 (5.23) 有多个几何约束 $x \in \Omega_i$, $i = 1,\cdots,n$, 显然它们可以简化为由交集给出的约束 $x \in \Omega := \bigcap\limits_{i=1}^{n} \Omega_i$. 于是可以利用如 5.1.1 小节中的关于基本法向量的交集法则来处理这些约束, 从而把 5.1.2~5.1.4 小节中相应的必要最优性和次优性条件推广到多个几何约束的情形. 为把 Fréchet 法向量表示的必要最优性和次优性条件推广到多个几何约束的情形, 与定理 5.21(i) 和定理 5.29 中所用的方法类似, 可以应用在 3.1.1 小节中所讨论的 Fréchet 法向量的模糊交法则.

还注意到, 除了通过利用下次微分变分原理得到下次优性条件外, 还可由定理 2.30 中的上次微分变分原理得到上次优性条件; 这方面的各种结果和讨论参见 Mordukhovich, Nam 与 Yen 的文章 [938].

5.2 具有均衡约束的数学规划

本节讨论一类特殊的最优化问题, 称为"均衡约束数学规划"(MPEC). 这类问题的典型特征是在约束中存在 $y \in S(x)$ 型的"均衡约束", 这里 $S(x)$ 通常表示一个"低层"参数最优化问题的解映射. MPEC 自然地出现在分层最优化和均衡理论的各个方面, 以及许多实际应用中, 尤其在与力学和经济学模型有关的应用中. 关于在有限维空间中这些问题的系统研究、例子和应用, 建议读者参阅专著 Luo, Pang 与 Ralph[820], Outrata, Kočvara 与 Zowe [1031] 及 Facchinei 与 Pang[424].

MPEC 中典型的均衡约束是不同类型的参数变分不等式、互补问题的解映射. 作为这个活跃的研究和应用的一个起点, 包含双层规划问题 (可追溯到 Stackelberg 对策论) 的一类 MPEC 是很重要的, 其中 $S(x)$ 是一参数线性或非线性规划问题的解映射. 值得注意的是, 几乎所有的 MPEC, 即使在具有互补约束的数学规划这样相对简单的情形中, 都本质地不同于具有等式和不等式约束的标准的非线性规划问题; 可能的简约导致各种不规则性, 例如, 导致 Mangasarian-Fromovitz 约束规范及类似的约束规范不成立.

在 5.2.1 小节中首先考虑的一般类 MPEC 具有以下以抽象形式:

$$\min \varphi(x,y) \quad \text{s.t.} \quad y \in S(x), \quad x \in \Omega, \tag{5.52}$$

这里, $S\colon X \rightrightarrows Y$ 是 Banach 空间之间的集值映射, $\varphi\colon X \times Y \to \overline{\mathbb{R}}, \Omega \subset X$. 这里的焦点是当均衡映射 S 以下列形式

$$S(x) := \{y \in Y \,|\, 0 \in f(x,y) + Q(x,y)\} \tag{5.53}$$

给出的情形, 其中 $f\colon X \times Y \to Z$, $Q\colon X \times Y \rightrightarrows Z$, 即 S 描述了第 4 章中研究的参数变分系统

$$0 \in f(x,y) + Q(x,y)$$

及其各种特殊形式的解映射. 我们知道, 模型 (5.53) 涵盖了经典变分不等式和互补问题以及 4.4 节中研究的它们的推广和修正问题的解映射. 下面将得到 (5.52), (5.53) 式给出的一般 MPEC 及其重要特殊情形的一阶必要最优性条件. 方法主要基于约化 MPEC 为 5.1 节中所考虑的具有几何约束的最优化问题, 同时考虑到它们的特殊结构, 然后利用 4.4 节建立的参数变分系统的灵敏性理论 (Lipschitz 稳定性的上导数估计和有效条件). 所得的结果涉及在复合形式下用来定义变分和半变分不等式的增广实值势函数的二阶次微分, 这对应用来说是最有意义的.

5.2.1 抽象 MPEC 的必要条件

本小节研究 (5.52) 型的抽象 MPEC, 给出下和上次微分形式的必要最优性条件. 这样的条件通过约化 (5.52) 为具有两个几何约束的标准形式 (5.1), 然后利用定理 5.5 的结果而得到. 这种方法可以利用 $X \times Y$ 上的乘积结构, 使得可以对 (5.52) 的初始数据给出适度的规范和 SNC 假设. 下面先从一般 MPEC 的上次微分必要最优性条件开始. 除非另行声明, 设 $\varphi\colon X \times Y \to \overline{\mathbb{R}}$ 为增广实值函数且在参考点处有限.

定理 5.33(抽象 MPEC 的上次微分最优性条件)　设 $(\bar{x}, \bar{y})$ 是 (5.52) 的局部最优解. 假设 X 和 Y 是 Asplund 空间, 集合 $\operatorname{gph} S$ 和 Ω 分别在 $(\bar{x}, \bar{y})$ 和 $\bar{x}$ 附近是局部闭的. 还假设 S 在 $(\bar{x}, \bar{y})$ 是 PSNC 的, 或者 Ω 在 $\bar{x}$ 是 SNC 的, 且混合规范条件

$$D_M^* S(\bar{x}, \bar{y})(0) \cap (-N(\bar{x}; \Omega)) = \{0\}$$

满足. 则

$$-x^* \in D_N^* S(\bar{x}, \bar{y})(y^*) + N(\bar{x}; \Omega)$$

对任意的 $(x^*, y^*) \in \widehat{\partial}^+ \varphi(\bar{x}, \bar{y})$ 成立.

证明　注意到 $(\bar{x}, \bar{y})$ 给出了函数 φ 受约束于 Asplund 空间 $X \times Y$ 中的约束 $(x, y) \in \Omega_1 := \operatorname{gph} S$ 和 $(x, y) \in \Omega_2 := \Omega \times Y$ 的一个局部极小值点. 对上面的问题应用定理 5.5(i) 中的上次微分条件, 易见 Ω_1 在 $(\bar{x}, \bar{y})$ 相对于 X 的 PSNC 性质简化为映射 S 在该点的 PSNC 性质, Ω_2 在 $(\bar{x}, \bar{y})$ 相对于 Y 总是强 PSNC 的, 在该点也是 SNC 的当且仅当 Ω 在 $\bar{x}$ 是 SNC 的. 而且, 定理的混合规范条件显然表明集合系统 $\{\Omega_1, \Omega_2\}$ 在 $(\bar{x}, \bar{y})$ 满足定义 3.2 意义下的极限规范条件. 因此, 根据定理 5.5(i) 有

$$-\widehat{\partial}^+ \varphi(\bar{x}, \bar{y}) \subset N\big((\bar{x}, \bar{y}); \operatorname{gph} S\big) + N(\bar{x}; \Omega) \times \{0\},$$

上式当然蕴涵着此定理的上次微分条件.　△

定理 5.33 的上次微分条件当 $\widehat{\partial}^+ \varphi(\bar{x}, \bar{y}) \neq \varnothing$ 时带有 MPEC 的非平凡的信息. 在 5.1.1 小节已讨论过满足这个要求的一些重要函数类. 回顾当 φ 在 $(\bar{x}, \bar{y})$ 是 Fréchet 可微的, 或在 $(\bar{x}, \bar{y})$ 附近是凹的和连续的, 则有 $\widehat{\partial}^+ \varphi(\bar{x}, \bar{y}) \neq \varnothing$. 如果 X 和 Y 都是 Asplund 空间, 后一种情形可以推广到在 $(\bar{x}, \bar{y})$ 附近 Lipschitz 连续和在该点上正则的函数类, 特别地, 这包括半凹函数类.

下面关于 MPEC 问题 (5.52) 中的局部极小值的更常见的下次微分条件与上面的上次微分条件相比本质是不同的, 一般来说它们是各自独立的; 参见注 5.4 中的讨论.

定理 5.34(抽象 MPEC 的下次微分最优性条件)　设 $(\bar{x}, \bar{y})$ 是 (5.52) 的局部最优解, X 和 Y 是 Asplund 空间, φ 在 $(\bar{x}, \bar{y})$ 附近是 l.s.c. 的, $\operatorname{gph} S$ 和 Ω 分别在

$(\bar{x},\bar{y})$ 和 $\bar{x}$ 附近是局部闭的. 则下列结论成立:

(i) 除了定理 5.33 的假设外, 还假设 φ 在 $(\bar{x},\bar{y})$ 是 SNEC 的, 条件

$$(x_\infty^*, y_\infty^*) \in \partial^\infty \varphi(\bar{x},\bar{y}), \quad 0 \in x_\infty^* + D_N^* S(\bar{x},\bar{y})(y_\infty^*) + N(\bar{x};\Omega)$$

只对 $x_\infty^* = y_\infty^* = 0$ 成立; 如果 φ 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 那么这些额外的假设自动成立. 则存在 $(x^*, y^*) \in \partial\varphi(\bar{x},\bar{y})$, 使得

$$0 \in x^* + D_N^* S(\bar{x},\bar{y})(y^*) + N(\bar{x};\Omega). \tag{5.54}$$

(ii) 假设 S 和 Ω 分别在 $(\bar{x},\bar{y})$ 和 $\bar{x}$ 附近是 SNC 的, 且规范条件

$$\Big[(x_\infty^*, y_\infty^*) \in \partial^\infty \varphi(\bar{x},\bar{y}), \quad x_1^* \in D_N^* S(\bar{x},\bar{y})(y_\infty^*), \quad x_2^* \in N(\bar{x};\Omega),$$
$$x_\infty^* + x_1^* + x_2^* = 0\Big] \Rightarrow x_\infty^* = y_\infty^* = x_1^* = x_2^* = 0$$

成立. 则存在 $(x^*, y^*) \in \partial\varphi(\bar{x},\bar{y})$ 使得最优性条件 (5.54) 成立.

证明 如同定理 5.33 的证明, 约化 (5.52) 为极小化 $\varphi(x,y)$, 受约束于几何约束 $(x,y) \in \Omega_1 = \operatorname{gph} S$ 和 $(x,y) \in \Omega_2 = \Omega \times Y$. 对这个问题应用定理 5.5, 易验证规范条件 (5.7) 化为定理在 (i) 中所假设的规范条件, 且下次微分最优性条件 (5.7) 给出了 (5.54). 类似地, 规范条件 (5.8) 简化为 (ii) 中所假设的规范条件, 从而完成了证明. △

基于类 Lipschitz 性质的上导数刻画, 下面建立定理 5.33 和定理 5.34 的有效推论, 通过关于所考虑的 MPEC 问题 (5.52) 的初始数据的内在要求, 保证上面的上和下次微分最优性条件都成立.

推论 5.35(类 Lipschitz 均衡约束下的上和下次微分条件) 设 $(\bar{x},\bar{y})$ 是 Asplund 空间 X 和 Y 中的 MPEC 问题 (5.52) 的局部最优解, 且 $\operatorname{gph} S$ 和 Ω 都是局部闭集. 假设 S 在 $(\bar{x},\bar{y})$ 附近是类 Lipschitz 的. 则有

$$-x^* \in D_N^* S(\bar{x},\bar{y})(y^*) + N(\bar{x};\Omega), \quad \forall (x^*, y^*) \in \widehat{\partial}^+ \varphi(\bar{x},\bar{y}).$$

进一步, 如果 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 的, 则

$$-x^* \in D_N^* S(\bar{x},\bar{y})(y^*) + N(\bar{x};\Omega)$$

对某 $(x^*, y^*) \in \partial\varphi(\bar{x},\bar{y})$ 成立.

证明 对 Asplund 空间之间的闭图映射 S 来说, 由定理 4.10 知 S 在 $(\bar{x},\bar{y})$ 附近是类 Lipschitz 的当且仅当 $D_M^* S(\bar{x},\bar{y})(0) = \{0\}$ 和 S 在 $(\bar{x},\bar{y})$ 是 PSNC 的. 因此 S 的类 Lipschitz 性质蕴涵着定理 5.33 中的所有假设, 从而保证上次微分最优性条

件成立. 另外, 如果 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则 $\partial^\infty\varphi(\bar{x},\bar{y})=\{0\}$, 因此根据定理 5.34(i), 有所述的下次微分最优性条件成立. △

根据推论 5.35, 对一般 MPEC 来说均衡约束的类 Lipschitz 性质是一个约束规范, 保证上和下两种次微分条件的标准形式成立. 现在, 如果 S 以 4.3 和 4.4 节中所考虑的约束和/或变分系统的参数形式给出, 利用这些节的结果可得到具有相应均衡约束 $y\in S(x)$ 的问题 (5.52) 的标准 (Karush-Kuhn-Tucker) 型的必要最优性条件, 它们给出了 $D_N^*S(\bar{x},\bar{y})$ 的上估计和这样映射的类 Lipschitz 性质的有效条件.

这里不使用 4.3 节中关于具有形式

$$S(x):=\left\{y\in Y\;\middle|\;g(x,y)\in\Theta,(x,y)\in\Omega\right\}$$

的参数约束系统的结果, 由于这种形式的约束对 MPEC 来说是不够具体的, 用这种方法所得的 (5.52) 的必要最优性条件与在 5.1.2 和 5.1.3 小节中所得到的关于具有算子和泛函约束问题的相应结果相比实际上没有带来新的信息. 下面把主要精力放在具有 4.4 节中所考虑的参数变分系统 (5.53) 控制的均衡约束的 (5.52) 型 MPEC 的必要最优性条件.

在建立这样的条件之前, 先对 (5.52) 型的抽象 MPEC 建立不强加任何约束规范的非规范 (Fritz John) 形式的一般必要最优性条件并结束本小节.

定理 5.36(非规范的 MPEC 的上和下次微分条件)　设 $(\bar{x},\bar{y})$ 是问题 (5.52) 的一个局部最优解. 假设 X 和 Y 是 Asplund 空间, 集合 $\operatorname{gph}S$ 和 Ω 分别在 $(\bar{x},\bar{y})$ 和 $\bar{x}$ 附近是局部闭的. 则在 S 在 $(\bar{x},\bar{y})$ 是 PSNC 的或 Ω 在 $\bar{x}$ 是 SNC 的条件下, 存在 $\lambda\in\{0,1\}$, 使得对任意的 $(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $x_1^*\in D_N^*S(\bar{x},\bar{y})(\lambda y^*)$ 和 $x_2^*\in N(\bar{x};\Omega)$ 满足

$$\lambda x^*+x_1^*+x_2^*=0,\quad (\lambda,x_1^*)\neq 0. \tag{5.55}$$

如果还假设 φ 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 则存在 $\lambda\geqslant 0,(x^*,y^*)\in\partial\varphi(\bar{x},\bar{y}),x_1^*\in D_N^*S(\bar{x},\bar{y})(\lambda y^*)$ 和 $x_2^*\in N(\bar{x};\Omega)$ 满足 (5.55).

证明　这个结果由定理 5.33 和定理 5.34 可得. 首先来证上次微分条件成立. 如果定理 5.33 中的混合规范条件满足, 则当 $\lambda=1$ 时在所给的假设下利用定理 5.33 中的断言有 (5.55) 成立. 如果该规范条件不成立, 则存在 $x^*\neq 0$ 满足

$$x^*\in D_M^*S(\bar{x},\bar{y})(0)\subset D_N^*S(\bar{x},\bar{y})(0)\quad 和\quad -x^*\in N(\bar{x};\Omega).$$

因此得 (5.55), 其中 $\lambda=0,x_1^*=x^*\neq 0$, 且 $x_2^*=-x^*$.

进一步, 假设 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 并且应用定理 5.34 中的断言 (i). 如果混合规范条件

$$D_M^*S(\bar{x},\bar{y})(0)\cap(-N(\bar{x};\Omega))=\{0\}$$

和定理 5.34(i) 中的其他假设被满足, 则由 5.34(i) 中的断言有关系 (5.55) 对 $\lambda = 1$ 和某 $(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})$ 成立. 如果这个混合规范条件不成立, 与上面上次微分条件的证明相同, 有 (5.55) 对 $\lambda = 0$ 和 $x_1^* \neq 0$ 成立. △

5.2.2 作为均衡约束的变分系统

本小节考虑均衡约束由依赖于参数的广义方程定义的 MPEC:

$$\min \varphi(x, y) \quad \text{s.t.} \quad 0 \in f(x, y) + Q(x, y), \quad x \in \Omega, \tag{5.56}$$

这里, $f\colon X \times Y \to Z$, $Q\colon X \times Y \rightrightarrows Z$ 分别是 Banach(大部分时间是 Asplund) 空间之间的单值和集值映射. 换言之, 模型 (5.56) 描述了由参数变分系统 $S(x)$ 控制的 (5.52) 型的 MPEC, 而 $S(x)$ 是扰动广义方程的解映射 (5.53). 下面的目标是根据问题 (5.56) 的初始数据 (φ, f, Q, Ω), 得到它的局部解的必要最优性条件. 这里基于定理 5.33 和定理 5.34 的结果得到 (5.56) 的上和下两种次微分最优性条件, 即得到标准/规范形式的必要条件. 类似定理 5.36, 也可由对应的非规范结果建立非标准形式的条件, 即与费用函数对应的乘子可以为零.

为从上面的定理中得到想要的必要条件, 需通过原始数据 f 和 Q 来表达涉及 (5.56) 中的均衡约束

$$y \in S(x) \Leftrightarrow 0 \in f(x, y) + Q(x, y)$$

的这些定理的假设和结论. 这可通过利用 4.4 节的结果来实现, 这些结果给出了这种映射 (变分系统)S 的上导数的上估计及其 PSNC 和类 Liplchitz 性质的充分条件. 具体来说, 针对 MPEC 的必要最优性条件中的应用, 由 4.4 中的结果实际上能得到的是通过 f 和 Q 表达的对 $D_N^* S(\bar{x}, \bar{y})$ 的上估计和关于 S 的 SNC 性质的充分条件. 由此得到保证约束规范

$$D_N^* S(\bar{x}, \bar{y})(0) \cap \big(-N(\bar{x}; \Omega) \big) = \{0\} \tag{5.57}$$

和定理 5.33 及定理 5.34 中的其他假设和结论成立的以 MPEC 问题 (5.56) 的初始数据和具体描述给出的充分条件.

下面首先给出 MPEC 问题 (5.56) 的上次微分必要最优性条件. 定理 5.37 给出了由 (5.56) 中一般参数变分系统控制的均衡约束情形的必要条件.

定理 5.37(具有一般变分约束的 MPEC 的上次微分条件) 设 $(\bar{x}, \bar{y})$ 是 MPEC 问题 (5.56) 的局部最优解, 这里 $f\colon X \times Y \to Z$, $Q\colon X \times Y \rightrightarrows Z$ 是 Asplund 空间之间的映射. 假设 f 在 $(\bar{x}, \bar{y})$ 附近是连续的, Ω 在 $\bar{x}$ 附近是局部闭的, Q 的图在 $(\bar{x}, \bar{y}, \bar{z})$ 附近是局部闭的, 其中 $\bar{z} := -f(\bar{x}, \bar{y})$. 还假设下面条件 (a)~(c) 其中之一成立:

(a) Ω 和 Q 分别在 $\bar{x}$ 和 $(\bar{x},\bar{y},\bar{z})$ 是 SNC 的, 并且下面两个规范条件满足:

$$\Big[(x^*,0)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*),-x^*\in N(\bar{x};\Omega)\Big]\Rightarrow x^*=0,$$

$$\Big[(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)\cap\big(-D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\big)\Big]\Rightarrow x^*=y^*=z^*=0;$$

当 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 第二个规范条件等价于

$$\Big[0\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\Big]\Rightarrow z^*=0. \tag{5.58}$$

(b) Ω 在 $\bar{x}$ 是 SNC 的, $\dim Z<\infty$, f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 并且规范条件

$$\Big[(x^*,0)\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*),-x^*\in N(\bar{x};\Omega)\Big]\Rightarrow x^*=0$$

和 (5.58) 成立.

(c) Q 在 $(\bar{x},\bar{y},\bar{z})$ 点是 SNC 的, f 在 $(\bar{x},\bar{y})$ 点是 PSNC(当 f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则它在该点自动是 PSNC) 的, 并且 (a) 中的规范条件都成立.

则对任意的 $(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $\widetilde{x}^*\in N(\bar{x};\Omega)$ 和 $z^*\in Z^*$, 满足必要最优性条件

$$(-x^*,-\widetilde{x}^*,-y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*).$$

证明 这里应用定理 5.33 中的上次微分最优性条件于问题 (5.56), 即当均衡约束 $y\in S(x)$ 以变分/广义方程形式 (5.53) 给出时的情形. 易见对 f 和 Q 所加的连续性和闭性假设保证 S 的局部闭性.

首先假设 Ω 在 $\bar{x}$ 是 SNC 的, 并且利用在定理 4.46 中所得到的对映射 $S(\cdot)$ 的上导数上估计. 于是有

$$D_N^*S(\bar{x},\bar{y})(y^*)\subset\big\{x^*\in X^*\big|\,\exists z^*\in Z^*\ \text{满足}$$
$$(x^*,-y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\big\}.$$

把此式代入规范条件 (5.57) 和定理 5.33 中的上次微分必要条件中, 可得到在 (a) 和 (b) 假设下此定理的结论.

现在考虑余下的情形, 即当 S 在定理 5.33 中是 PSNC 的情形, 并且根据 f 和 Q 给出有效条件保证 S 的该性质 (甚至 SNC) 成立. 这实际上在定理 4.59 的证明中已经做过了, 即基于定理 3.84 中 SNC 分析法则来验证 S 的类 Lipschitz 性质的上导数准则那部分. 利用这些结果, 在 (c) 的假设下得到定理的上次微分最优性条件. △

接下来基于定理 5.34 得到 MPEC (5.56) 的下次微分最优性条件, 其中用定理 4.46 和定理 3.84 中的结果来处理 (5.56) 中的均衡约束 S.

定理 5.38(具有一般变分约束的 MPEC 的下次微分条件) 设 $(\bar{x},\bar{y})$ 是 (5.56) 的局部最优解, 其中 $f\colon X\times Y\to Z$, $Q\colon X\times Y\rightrightarrows Z$ 是 Asplund 空间之间的映射. 假设 φ 在 $(\bar{x},\bar{y})$ 附近是 l.s.c. 的, f 在 $(\bar{x},\bar{y})$ 附近是连续的, Ω 在 $\bar{x}$ 附近是局部闭的, Q 的图在 $(\bar{x},\bar{y},\bar{z})$ 附近是局部闭的, 其中 $\bar{z}=-f(\bar{x},\bar{y})$. 则下面的断言成立:

(i) 除了定理 5.37 中的假设, 进一步设函数 φ 在 $(\bar{x},\bar{y})$ 点 SNEC, 假设条件

$$(x_\infty^*-\widehat{x}^*,-y_\infty^*)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*),$$

$$(x_\infty^*,y_\infty^*)\in\partial^\infty\varphi(\bar{x},\bar{y})\quad (\text{对某个}\ \ \widehat{x}^*\in N(\bar{x};\Omega),z^*\in Z^*\ \ \text{成立})$$

只当 $x_\infty^*=y_\infty^*=0$ 时成立; 如果 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则上述两个条件都自动成立. 则存在 $(x^*,y^*)\in\partial\varphi(\bar{x},\bar{y}),\widetilde{x}^*\in N(\bar{x};\Omega),z^*\in Z^*$ 满足

$$(-x^*-\widetilde{x}^*,-y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*). \tag{5.59}$$

(ii) 假设 Ω 和 Q 分别在 $\bar{x}$ 和 $(\bar{x},\bar{y},\bar{z})$ 是 SNC 的, f 在 $(\bar{x},\bar{y})$ 是 PSNC 的, 规范条件

$$\Big[(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)\cap\big(-D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\big)\Big]\Rightarrow x^*=y^*=z^*=0,$$

$$\Big[(x_\infty^*,y_\infty^*)\in\partial^\infty\varphi(\bar{x},\bar{y}),(x_1^*,-y_\infty^*)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*),$$

$$x_2^*\in N(\bar{x};\Omega),z^*\in Z^*,x_\infty^*+x_1^*+x_2^*=0\Big]\Rightarrow x_\infty^*=y_\infty^*=x_1^*=x_2^*=0$$

成立. 则存在 $(x^*,y^*)\in\partial\varphi(\bar{x},\bar{y})$, $\widetilde{x}^*\in N(\bar{x};\Omega)$, $z^*\in Z^*$ 满足最优性条件 (5.59).

证明 为证 (i) 成立, 应用定理 5.34(i), 然后分别利用定理 4.46 和定理 4.59 中所得的均衡映射 S 的上导数上估计和 SNC 性质的有效条件, 类似定理 5.37 的证明可得结论成立. 断言 (ii) 基于定理 5.34(ii) 利用同样的方法可证. △

下面给出定理 5.37 和定理 5.38(i) 的有效推论保证 (5.56) 中均衡映射 S 的类 Lipschitz 性质, 从而定理 5.33 和 5.34(i) 中的规范条件和 PSNC 条件成立.

推论 5.39(类 Lipschitz 变分约束下的上和下次微分条件) 设 $(\bar{x},\bar{y})$ 是 (5.56) 的局部最优解, 其中 $f\colon X\times Y\to Z$, $Q\colon X\times Y\rightrightarrows Z$ 是 Asplund 空间之间的映射, 假设 f 在 $(\bar{x},\bar{y})$ 附近是连续的, Ω 在 $\bar{x}$ 附近是局部闭的, Q 的图在 $(\bar{x},\bar{y},\bar{z})$ 附近是局部闭的, 其中 $\bar{z}=-f(\bar{x},\bar{y})$, 假设规范条件

$$\Big[(x^*,0)\in D_N^*f(\bar{x},\bar{y})(z^*)+D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\Big]\Rightarrow x^*=0,$$

$$\Big[(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)\cap\big(-D_N^*Q(\bar{x},\bar{y},\bar{z})(z^*)\big)\Big]\Rightarrow x^*=y^*=z^*=0$$

满足. 则对任意的 $(x^*, y^*) \in \widehat{\partial}^+\varphi(\bar{x}, \bar{y})$, 都存在 $\widetilde{x}^* \in N(\bar{x}; \Omega)$ 和 $z^* \in Z^*$, 使得最优性条件 (5.59) 成立. 进一步, 如果 φ 在 $(\bar{x}, \bar{y})$ 附近是 Lipschitz 连续的, 则存在 $(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y}), \widetilde{x}^* \in N(\bar{x}; \Omega)$ 和 $z^* \in Z^*$ 满足 (5.59).

证明 这些结论分别由定理 5.37 和 5.38(i) 直接可得. 它们也是推论 5.35 和定理 4.59 的推论, 保证了 MPEC 问题 (5.56) 中的均衡约束 S 的类 Lipschitz 性质成立. △

利用 f 和/或 Q 的上导数表示, 在某些特殊的情形下易得所得结果的具体和简化形式; 特别地, 可对比 4.4 节中的严格可微映射 f 和凸图多值函数 Q 以及不依赖参数的域 $Q = Q(y)$.

下面更详细地讨论 (5.56) 中的变分约束最令人感兴趣的情形, 即当均衡映射 S 以次微分形式给出时的情形, 它包括经典变分不等式、互补问题和半变分不等式以及它们的进一步推广. 这里主要关注 4.4 节中讨论的具有复合次微分结构的两类广义变分不等式 (GVI), 其中均衡映射 S 以 (4.66) 和 (4.67) 形式给出. 第一类 GVI 诱导的 MPEC 类型为

$$\min \varphi(x, y) \quad \text{s.t.} \quad 0 \in f(x, y) + \partial(\psi \circ g)(x, y), \quad x \in \Omega, \tag{5.60}$$

由 Banach 空间之间的单值映射 $f \colon X \times Y \to X^* \times Y^*$, $g \colon X \times Y \to W$ 和增广实值函数 $\psi : W \to \overline{\mathbb{R}}$ 控制, 得到关于 (5.60) 的上和下两种次微分必要最优性条件. 为简单起见, 对下次微分条件的情形只考虑局部 Lipschitz 费用函数 φ. 这里从 (5.60) 中的映射 $g \colon Y \to W$ 是光滑 (导数是满射) 和参数独立的映射情形开始. 这使得在用 $\varphi \colon W \to \overline{\mathbb{R}}$ 的 (基本的) 二阶次微分表达的关于 (5.60) 的必要最优性条件中允许空间的一般性. 沿用 4.4 节的术语, 称这样的问题为由带有复合势函数的参数半变分不等式 (HVI) 控制的 MPEC.

定理 5.40(由带有复合势函数的 (HVI) 控制的 MPEC 的上、下次微分条件) 设 $(\bar{x}, \bar{y})$ 是 (5.60) 的局部最优解, 其中 $f \colon X \times Y \to Y^*$, $g \colon Y \to W$, $\psi \colon W \to \overline{\mathbb{R}}$. 假设 W 是 Banach 空间, X 是 Asplund 空间, Y 是有限维空间, 还假设下面的条件成立:

(a) f 在 $(\bar{x}, \bar{y})$ 是严格可微的, 具有满射偏导数 $\nabla_x f(\bar{x}, \bar{y}) : X \to Y^*$.

(b) g 在 $\bar{y}$ 附近是 $\mathcal{C}^1$ 的, 具有满射导数 $\nabla g(\bar{y}) \colon Y \to W$, 映射 $\nabla g \colon Y \to \mathcal{L}(Y, W)$ 在 $\bar{y}$ 是严格可微的.

(c) Ω 在 $\bar{x}$ 附近是局部闭的, $\partial\psi$ 的图在 $(\bar{w}, \bar{v})$ 附近是局部闭的, 其中 $\bar{w} := g(\bar{y}), \bar{v} \in W^*$ 是满足关系

$$-f(\bar{x}, \bar{y}) = \nabla g(\bar{y})^* \bar{v}, \quad \bar{v} \in \partial\psi(\bar{w})$$

的唯一泛函. 如果 $u=0$ 是满足包含系统

$$\begin{cases} 0 \in \nabla_x f(\bar{x},\bar{y})^* u + N(\bar{x};\Omega), \\ 0 \in \nabla_y f(\bar{x},\bar{y})^* u + \nabla^2\langle \bar{v}, g\rangle(\bar{y})^* u + \nabla g(\bar{y})^* \partial_N^2 \psi(\bar{w},\bar{v})(\nabla g(\bar{y})u) \end{cases}$$

的唯一向量, 则对任意的 $(x^*,y^*) \in \widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $u \in Y$ 使得

$$\begin{aligned} -x^* &\in \nabla_x f(\bar{x},\bar{y})^* u + N(\bar{x};\Omega), \\ -y^* &\in \nabla_y f(\bar{x},\bar{y})^* u + \nabla^2\langle \bar{v}, g\rangle(\bar{y})^* u + \nabla g(\bar{y})^* \partial_N^2 \psi(\bar{w},\bar{v})\big(\nabla g(\bar{y})u\big). \end{aligned} \tag{5.61}$$

进一步, 如果 φ 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 则存在 $u \in Y, (x^*,y^*) \in \partial\varphi(\bar{x},\bar{y})$ 满足 (5.61).

证明 为建立定理的上次微分条件, 当 $Q(y) := \partial(\psi \circ g)(y)$ 时在假设 (c) 下利用定理 5.37 的结果. 考虑到 f 在 $(\bar{x},\bar{y})$ 的严格可微性且具有满射导数 $\nabla_x f(\bar{x},\bar{y})$, 以及 Q 的参数独立性, 则条件 (5.58) 自动成立, 同时定理 5.37(a) 中的第一个规范条件在映射 $\partial(\psi \circ g)(y)(\cdot)$ 在 $(\bar{y},\bar{z})$ 附近是局部闭图的条件下简化为

$$\Big[0 \in \nabla_x f(\bar{x},\bar{y})^* u + N(\bar{x};\Omega), 0 \in \nabla_y f(\bar{x},\bar{y})^* u + \partial^2(\psi \circ g)(\bar{y},\bar{z})(u)\Big] \Rightarrow u = 0,$$

其中 $\bar{z} := -f(\bar{x},\bar{y})$. 注意到 Q 的 SNC 性质和 f 在参考点的 PSNC 性质由 Y 的有限维性和 f 的严格可微性立即可得. 因此, 通过把定理 5.37 中的上次微分最优性条件应用于 (5.60), 对任意的上次梯度 $(x^*,y^*) \in \widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $u \in Y$ 使得

$$-x^* \in \nabla_x f(\bar{x},\bar{y})^* u + N(\bar{x};\Omega), \quad -y^* \in \nabla_y f(\bar{x},\bar{y})^* u + \partial^2(\psi \circ g)(\bar{y},\bar{z})(u).$$

现应用命题 1.112 (i) 中的一阶次微分链式法则, 有等式

$$\partial(\psi \circ g)(y) = \nabla g(y)^* \partial\psi(w)$$

对任意接近于 $\bar{y}$ 的 y 和 $w = g(y)$ 成立, 这蕴涵着 $\partial(\psi \circ g)(\cdot)$ 的图在 $(\bar{y},\bar{z})$ 附近是局部闭的当且仅当次微分映射 $\partial\psi(\cdot)$ 在 $(\bar{w},\bar{v})$ 附近是闭图的. 进一步对 $\partial^2(\psi \circ g)(\bar{y},\bar{z})$ 在上面等式中应用定理 1.127 的二阶次微分链式法则, 并且考虑到在所给的假设下有 $\nabla g(\bar{y})^{**} = \nabla g(\bar{y})$, 即得定理所陈述的上次微分条件.

如果 φ 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 则定理的下次微分利用类似方法由定理 5.38(i) 可以推得. △

回顾如果 ψ 在 $\bar{w},\bar{v}$ 附近是连续的或在该点是顺从的 (参见 3.2.5 小节中的定义), 则在上面定理假设 (c) 下 $\partial\psi$ 在 $(\bar{w},\bar{v})$ 附近自动是闭图的.

下一结果给出了有限维空间中带有复合势函数由依赖于参数的 GVI 控制的 MPEC 问题 (5.60) 的必要最优性条件; 与定理 5.40 中的假设相比, 对 f 和 g 的限制本质地减少了, 但对 ψ 的限制却增加了.

定理 5.41(带有复合势函数的 GVI 控制的 MPEC 的上、下次微分条件) 设 $(\bar{x},\bar{y})$ 是 (5.60) 的局部最优解, 其中 $f\colon X\times Y\to X^*\times Y^*$ 和 $g\colon X\times Y\to W$ 是有限维空间之间的映射. 假设 f 在 $(\bar{x},\bar{y})$ 附近是连续的, g 在该点附近是二次连续可微的, ψ 在 $\bar{w}:=g(\bar{x},\bar{y})$ 附近是 l.s.c.的, Ω 在 $\bar{x}$ 附近是局部闭的. 记 $\bar{z}:=-f(\bar{x},\bar{y})\in\partial(\psi\circ g)(\bar{x},\bar{y})$, $M(\bar{x},\bar{y}):=\{\bar{v}\in W^*|\bar{v}\in\partial\psi(\bar{w}),\nabla g(\bar{x},\bar{y})^*\bar{v}=\bar{z}\}$, 并假设:

(a) 函数 ψ 在 $\bar{w}$ 附近是下正则的, 当 w 在 $\bar{w}$ 附近时 $\partial\psi$ 和 $\partial^\infty\psi$ 的图是闭的.

(b) 下面的关于复合 $\psi\circ g$ 的一阶和二阶规范条件成立:

$$\partial^\infty\psi(\bar{w})\cap\ker\nabla g(\bar{x},\bar{y})^*=\{0\},$$

$$\partial^2\psi(\bar{w},\bar{v})(0)\cap\ker\nabla g(\bar{x},\bar{y})^*=\{0\},\quad\forall\,\bar{v}\in M(\bar{x},\bar{y}).$$

(c) 下述两个关系成立:

$$\begin{aligned}&\Big[(x^*,y^*)\in\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\big(\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})u+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\big)\\&\cap\big(-D^*f(\bar{x},\bar{y})(u)\big)\Big]\Rightarrow(x^*,y^*,u)=(0,0,0),\end{aligned}$$

$$\begin{aligned}&\Big[(x^*,0)\in D^*f(\bar{x},\bar{y})(u)+\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\big(\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})(u)\\&+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\big),-x^*\in N(\bar{x};\Omega)\Big]\Rightarrow x^*=0.\end{aligned}$$

则对任意的 $(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $u\in X\times Y$ 使得

$$\begin{aligned}(-x^*,-y^*)\in&D^*f(\bar{x},\bar{y})(u)+\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\Big[\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})(u)\\&+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\Big]+N(\bar{x};\Omega).\end{aligned}\tag{5.62}$$

更进一步, 如果 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则存在 $(x^*,y^*)\in\partial\varphi(\bar{x},\bar{y})$ 和 $u\in X\times Y$ 满足 (5.62).

证明 在

$$S(x):=\big\{y\in Y\,|\,0\in f(x,y)+\partial(\psi\circ g)(x,y)\big\}$$

情形下利用定理 5.33 和定理 5.34(i) 中的上和下次微分最优性条件, 再利用在二阶次微分和法则的基础上而得的定理 4.50 中的 $D_N^*S(\bar{x},\bar{y})$ 的上估计. 这样在所给的假设下就得到定理的结论. △

当 f 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 则上面定理中 (c) 中的第一个关系简化为

$$0\in\partial\langle u,f\rangle(\bar{x},\bar{y})+\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\Big[\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})^*u+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\Big]\Rightarrow u=0.$$

如果 $g=g(y), f$ 在 $(\bar{x},\bar{y})$ 是严格可微的且具有满射偏导数 $\nabla_x f(\bar{x},\bar{y})$, 则此关系自动成立.

如果 (5.60) 中的均衡约束的势函数 $\phi:=\psi\circ g$ 是强顺从的 (参见 3.25 小节中的定义), 则定理 5.41 中的一阶假设恰好自动满足.

推论 5.42(带有顺从势函数的 MPEC 的最优性条件) 设 $(\bar{x},\bar{y})$ 是有限维空间中 MPEC 问题 (5.60) 的局部最优解, 其中 Ω 在 $\bar{x}$ 附近是闭的, f 在 $(\bar{x},\bar{y})$ 附近连续, 势函数 $\phi=\psi\circ g$ 在该点是强顺从的. 假设定理 5.41(c) 中的假设和 (b) 中的二阶规范条件满足. 在没有其他假设下则有此定理的上次微分最优性条件. 进一步, 如果 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则上面的下次微分条件也成立.

证明 根据在 3.2.5 小节中所讨论的强顺从函数的性质和在推论 3.76 中所给的二阶次微分链式法则, 由定理 5.41 可得结论成立. △

现在考虑由具有复合域的广义变分不等式控制的 MPEC:

$$\min\varphi(x,y)\quad\text{s.t.}\quad 0\in f(x,y)+(\partial\psi\circ g)(x,y),\quad x\in\Omega.\tag{5.63}$$

其中 $g\colon X\times Y\to W$, $\psi\colon W\to\overline{\mathbb{R}}$, $f\colon X\times Y\to W^*$. 下面的定理给出了这样的 MPEC 的上和下次微分类型的一般必要最优性条件.

定理 5.43(由具有复合域 GVI 控制的 MPEC 的上、下次微分条件) 设 $(\bar{x},\bar{y})$ 是问题 (5.63) 的局部最优解, 其中 Ω 在 $\bar{x}$ 附近是闭的, $\bar{w}:=g(\bar{x},\bar{y})$, $\bar{z}:=-f(\bar{x},\bar{y})$. 则下面的断言成立:

(i) 假设 X,Y 是 Asplund 空间, 而 W 是 Banach 空间, $g=g(y)$ 在 $\bar{y}$ 是严格可微的, 且具有满射导数 $\nabla g(\bar{y})$, f 在 $(\bar{x},\bar{y})$ 是严格可微的, 且具有满射偏导数 $\nabla_x f(\bar{x},\bar{y})$, $u=0\in W^{**}$ 是满足

$$0\in\nabla_x f(\bar{x},\bar{y})^*u+N(\bar{x};\Omega),\quad 0\in\nabla_y f(\bar{x},\bar{y})^*u+\nabla g(\bar{y})^*\partial_N^2\psi(\bar{w},\bar{z})(u)$$

的唯一元素. 如果 Ω 在 $\bar{x}$ 是 SNC 的, 或 $\partial\psi$ 在 $(\bar{w},\bar{z})$ 是 SNC 的, 则对任意的 $(x^*,y^*)\in\widehat{\partial}^+\varphi(\bar{x},\bar{y})$, 存在 $u\in W^{**}$ 使得

$$\begin{aligned}&-x^*\in\nabla_x f(\bar{x},\bar{y})^*u+N(\bar{x};\Omega),\\&-y^*\in\nabla_y f(\bar{x},\bar{y})^*u+\nabla g(\bar{y})^*\partial_N^2\psi(\bar{w},\bar{z})(u).\end{aligned}\tag{5.64}$$

(ii) 假设 X, Y, W, W^* 是 Asplund 空间, f 和 g 在 $(\bar{x}, \bar{y})$ 附近是连续的, $\partial\psi$ 的图在 $(\bar{w}, \bar{z})$ 附近是范数闭的, 假设

$$\partial_N^2\psi(\bar{w}, \bar{z})(0) \cap \ker D_N^* g(\bar{x}, \bar{y}) = \{0\},$$

而且 $x^* = 0$ 是满足

$$(x^*, 0) \in D_N^* f(\bar{x}, \bar{y})(u) + D_N^* g(\bar{x}, \bar{y}) \circ \partial_N^2 \psi(\bar{w}, \bar{z})(u), \quad -x^* \in N(\bar{x}; \Omega)$$

对某 $u \in W^{**}$ 成立的唯一元素, $(x^*, y^*, u) = (0, 0, 0)$ 是满足

$$(x^*, y^*) \in D_N^* f(\bar{x}, \bar{y})(u) \cap (-D_N^* g(\bar{x}, \bar{y}) \circ \partial_N^2 \psi(\bar{w}, \bar{z})(u))$$

的唯一元素组. 如果 f 在 $(\bar{x}, \bar{y})$ 附近是 Lipschitz 连续的, 且 $\dim W < \infty$, 或者 g 在 $(\bar{x}, \bar{y})$ 是 PSNC 的, 且 $\partial\psi$ 在 $(\bar{z}, \bar{w})$ 是 SNC 的, 或者 g 在 $(\bar{x}, \bar{y})$ 是 SNC 的, 且 $\partial\psi^{-1}$ 在 $(\bar{z}, \bar{w})$ 是 PSNC 的, 则对任意的上次梯度 $(x^*, y^*) \in \widehat{\partial}^+ \varphi(\bar{x}, \bar{y})$, 存在 $\widetilde{x}^* \in N(\bar{x}; \Omega)$ 和 $u \in W^{**}$ 使得

$$(-x^* - \widetilde{x}^*, -y^*) \in D_N^* f(\bar{x}, \bar{y})(u) + D_N^* g(\bar{x}, \bar{y}) \circ \partial_N^2 \psi(\bar{w}, \bar{z})(u). \tag{5.65}$$

(iii) 假设 φ 在 $(\bar{x}, \bar{y})$ 附近是 Lipschitz 连续的, 而且 (i) 或 (ii) 中的假设成立. 则存在 $(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})$ 和 $u \in W^{**}$ 满足 (5.64), 或存在 $(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y})$, $\widetilde{x}^* \in N(\bar{x}; \Omega), u \in W^{**}$ 满足 (5.65).

证明 为建立 (i), 把定理 5.33 中的上次微分最优性条件和命题 4.53 中的上导数公式应用于 (5.63) 中的均衡映射

$$S(x) := \{y \in Y \,|\, 0 \in f(x, y) + (\partial\psi \circ g)(x, y)\}.$$

由定理 1.22 得, 在 $\nabla g(\bar{y})$ 的满射假设下, S 在 $(\bar{x}, \bar{y})$ 的 SNC 性质等价于 $\partial\psi$ 在 $(\bar{w}, \bar{z})$ 的 SNC 性质. 于是合并定理 5.33 和命题 4.53 的假设和结论, 即得 (i) 成立. 根据定理 5.33 的最优性条件和定理 4.54 的上导数上估计, (ii) 的证明是类似的. 正如定理 4.54 的证明, 复合 $\partial\psi \circ g$ 的 SNC 性质的充分条件由定理 3.98 可得.

(iii) 中的下次微分最优性条件通过利用上面的讨论由定理 5.34(i) 可得. △

若严格可微映射 f 和 g 的像空间是有限维的 (可能有不满射导数), 下面给出定理 5.43 的上和下次微分断言 (ii) 和 (iii) 的一些推论, 此时定理中的关系可以有本质的简化.

推论 5.44(由具有复合域 GVI 控制的特殊 MPEC 的最优性条件) 设 $(\bar{x}, \bar{y})$ 是问题 (5.63) 的局部最优解, 其中 $f\colon X \times Y \to \mathbb{R}^m$, $g\colon X \times Y \to \mathbb{R}^m$ 在 $(\bar{x}, \bar{y})$ 是

严格可微的, $\Omega \subset X$ 在 $\bar{x}$ 附近是闭的. 假设 X 和 Y 是 Asplund 空间, $\mathrm{gph}\,\partial\psi$ 在 $(\bar{w}, \bar{z})$ 附近是闭的 (对连续函数和顺从函数来说这个条件自动成立), 且假设

$$\partial^2\psi(\bar{w}, \bar{z})(0) \cap \ker\nabla g(\bar{x}, \bar{y})^* = \{0\}$$

成立, 包含系统

$$\begin{cases} x^* \in \nabla_x f(\bar{x}, \bar{y})^* u + \nabla_x g(\bar{x}, \bar{y})^* \partial^2\psi(\bar{w}, \bar{z})(u), \quad -x^* \in N(\bar{x}; \Omega), \\ 0 \in \nabla_y f(\bar{x}, \bar{y})^* u + \nabla_y g(\bar{x}, \bar{y})^* \partial^2\psi(\bar{w}, \bar{z})^*(u) \end{cases}$$

有唯一平凡解 $x^* = u = 0$. 则对任意的 $(x^*, y^*) \in \widehat{\partial}^+\varphi(\bar{x}, \bar{y})$, 存在 $u \in \mathbb{R}^m$ 使得

$$\begin{aligned} -x^* \in & \nabla_x f(\bar{x}, \bar{y})^* u + \nabla_x g(\bar{x}, \bar{y})^* \partial^2\psi(\bar{w}, \bar{z})(u) + N(\bar{x}; \Omega) \\ & - y^* \in \nabla_y f(\bar{x}, \bar{y})^* u + \nabla_y g(\bar{x}, \bar{y})^* \partial^2\psi(\bar{w}, \bar{z})^*(u). \end{aligned} \tag{5.66}$$

进一步, 如果费用函数 φ 在 $(\bar{x}, \bar{y})$ 附近是 Lipschitz 连续的, 则存在 $(x^*, y^*) \in \partial\varphi(\bar{x}, \bar{y}), u \in \mathbb{R}^m$ 满足 (5.66).

证明 根据严格可微函数的上导数表示, 由定理 5.43(ii, iii) 易得结论成立. △

注 5.45(正常扰动下 MPEC 的最优性条件) 考虑由正常扰动广义方程

$$\Sigma(x, z) := \left\{ y \in Y \middle| z \in f(x, y) + Q(x, y) \right\}$$

的解映射给出的均衡约束的一类 MPEC:

$$\min \quad \varphi(x, z, y) \quad \text{s.t.} \quad z \in f(x, y) + Q(x, y), \, (x, z) \in \Omega. \tag{5.67}$$

它可看成是关于参数对 $p := (x, z)$ 的 MPEC(5.56) 的特殊情形. 因此上面关于 (5.56) 所得的必要最优性条件即蕴涵着关于 (5.67) 的相应结果. 另一方面, (5.67) 中的参数依赖均衡约束的正常结构允许对这类 MPEC 得到特殊的结果. 在 4.4.3 小节中所得的均衡映射 $\Sigma(\cdot)$ 类 Lipschitz 性质的有效条件下, 以及在定理 5.33 和定理 5.34(也可参见推论 5.35) 的上和下次微分最优化条件的基础之上, 这一点是能够实现的. 这些特殊结果自动产生 MPEC(5.67) 以及 (5.56) 的必要最优性条件, 它们通常与在推论 5.39 中所得的结果及其具体化是独立的. 请参阅 4.4.3 小节中的相应结果和讨论.

5.2.3 利用精确惩罚的 MPEC 的修正下次微分条件

本小节建立了求由 (5.56) 型的参数变分系统控制的 MPEC 的必要最优性条件的另一方法. 与前一小节相比, 这种方法不直接对 5.2.1 小节中的一般最优性条件

应用分析法则, 而是在开始利用惩罚程序, 这在某些情形下产生一些精细的下次微分结果. 另一方面, 惩罚方法无法得到在 5.2.2 小节中给出的上次微分型的必要最优性条件.

首先定义集值映射在它们图像的参考点处的一种 Lipschitz 性质.

定义 5.46(集值映射的平静性) 设 $F: X \rightrightarrows Y$ 是 Banach 空间之间的集值映射, $(\bar{x},\bar{y}) \in \operatorname{gph} F$. 则称 F 在 $(\bar{x},\bar{y})$ 是平静的, 且具有模 $\ell \geqslant 0$, 如果存在 $\bar{x}$ 的邻域 U 和 $\bar{y}$ 的邻域 V, 使得

$$F(x) \cap V \subset F(\bar{x}) + \ell \parallel x - \bar{x} \parallel \mathbb{B}, \quad \forall x \in U. \tag{5.68}$$

如果在 (5.68) 中可选取 $V = Y$ 且 $\bar{x} \in \operatorname{dom} F$, 则称映射 F 在点 $\bar{x}$ 处是平静的.

集值映射在它们有效域内的点处的第二种平静性性质也称为 F 在 $\bar{x} \in \operatorname{dom} F$ 的上 Lipschitz 性质 (在 Robinson 意义下). 沿用本书的术语, 图局部化的平静性性质 (5.68) 似乎也可称为 F 在 $(\bar{x},\bar{y}) \in \operatorname{gph} F$ 的类上 Lipschitz 性质.

易见上面集值映射的平静性/上 Lipschitz 性质与定义 1.40 中"完全的"Lipschitz 性质相比有更少的限制, 那里 $\bar{x}$ 被 $u \in U$ 替代, 与 x 一起在 $\bar{x}$ 附近变化. 另一方面, 平静性性质与完全的 Lipschitz 性质相比, 其相对于参考点 $\bar{x}$ 的扰动不是鲁棒的. 而且, 上面的平静性性质不分别蕴涵着 $(\bar{x},\bar{y}) \in \operatorname{int}(\operatorname{gph} F)$ 和 $\bar{x} \in (\operatorname{dom} F)$. 还注意到对单值映射 $F = f: X \to Y$ 来说 f 的平静性性质不能简化为单值映射的标准局部 Lipschitz 性质. Robinson 考虑过一种经典的情形, 映射 $F: \mathbb{R}^n \rightrightarrows \mathbb{R}^m$ 在任意一点 $\bar{x} \in \operatorname{dom} F$ 处是平静的/上 Lipschitz 的, 但在 $\bar{x}$ 附近可能不是局部 Lipschitz 的, 即当 F 是分片多面体映射, 它的图可表示为有限多 (凸) 多面体集的并集. 这样的映射对在具有有限多等式和不等式型线性约束的数学规划中的应用来说非常重要.

这一小节, 利用平静性性质 (5.68) 研究由参数变分系统控制的 MPEC. 首先考虑下面含有非参数广义方程类型约束的最优化问题:

$$\min \varphi(t) \quad \text{s.t.} \quad 0 \in F(t), \quad t \in \Omega, \tag{5.69}$$

其中 $F: T \rightrightarrows Z$ 是 Banach 空间之间的集值映射, $\varphi: T \to \overline{\mathbb{R}}, \Omega \subset T$. 由于 (5.69) 中的约束可以记为 $t \in F^{-1}(0) \cap \Omega$, 故这个问题是 5.1.2 小节中最优化问题 (5.12) 的一种特殊情形. 应用在那里得到的关于问题 (5.12) 的必要最优性条件时, 由于 $\Theta = \{0\}$ 时的规范条件 (5.15), 就不可避免地要求 F^{-1} 在最小值点处附近的类 Lipschitz 性质（或 F 的度量正则性性质）. 然而, 通过利用初步的一个精确惩罚程序, 这个要求可以得到放松.

事实上, 问题 (5.69) 可等价地写为

$$\min \varphi(x) \quad \text{s.t.} \quad z \in F(t), \quad z = 0, \quad t \in \Omega.$$

下面的辅助结果, 紧密地相关于定理 5.16, 把由 (5.69) 化归为 5.2.1 小节中一般 MPEC 问题.

引理 5.47(广义方程约束下的精确惩罚) 设 $\bar{t}$ 是 Banach 空间框架下的问题 (5.69) 的一个局部最优解. 假设 φ 在 $\bar{t}$ 附近是 Lipschitz 连续的, 且具有模 ℓ_φ, 映射 $(F^{-1}\cap\Omega)(z):=F^{-1}(z)\cap\Omega$ 在 $(0,\bar{t})$ 处是平静的, 且具有模 ℓ. 如果 $\mu\geqslant\ell_\varphi\cdot\ell$, 则存在 $\bar{t}$ 的邻域 V, $0\in Z$ 的邻域 U 使得 $(\bar{t},0)\in T\times Z$ 是惩罚问题

$$\min\quad \psi(t,z):=\varphi(t)+\mu\|z\|\quad \text{s.t.}\quad z\in F(t)\cap U,\ t\in\Omega\cap V$$

的解.

证明 由于 $F^{-1}\cap\Omega$ 在 $(0,\bar{t})$ 处是平静的, 且具有模 $\ell\geqslant 0$, 故存在 $\bar{t}$ 的邻域 V, $0\in Z$ 的邻域 U, 使得对某个 $\widehat{t}\in F^{-1}\cap\Omega$, 有估计

$$\|t-\widehat{t}\|\leqslant\ell\|z\|,\quad \forall t\in F^{-1}(z)\cap\Omega\cap V,\quad z\in U.$$

利用这个估计和具有模 ℓ_φ 的 φ 的 Lipschitz 连续性, 有

$$\begin{aligned}\varphi(\bar{t})&\leqslant\varphi(\widehat{t})=\varphi(t)+(\varphi(\widehat{t})-\varphi(t))\\&\leqslant\varphi(t)+\ell_\varphi\|\widehat{t}-t\|\leqslant\varphi(t)+\ell_\varphi\cdot l\|z\|\\&\leqslant\varphi(t)+\mu\|z\|\end{aligned}$$

对任意的 $t\in F^{-1}(z)\cap\Omega\cap V$, $z\in U$ 和 $\mu\geqslant\ell_\varphi\cdot\ell$ 成立. △

定理 5.48(广义方程约束下必要最优性条件) 设 $\bar{t}$ 是问题 (5.69) 的局部最优解, 其中 T 和 Z 是 Asplund 空间, Ω 和 $\operatorname{gph}F$ 分别在 $\bar{t}$ 和 $(\bar{t},0)$ 附近是局部闭的. 假设 φ 在 $\bar{t}$ 附近是局部 Lipschitz 的, 且具有模 ℓ_φ, $F^{-1}\cap\Omega$ 在 $(0,\bar{t})$ 是平静的, 且具有模 ℓ, 假设混合规范条件

$$D^*_MF(\bar{t},0)(0)\cap(-N(\bar{t};\Omega))=\{0\}$$

成立. 还假设 F 在 $(\bar{t},0)$ 是 PSNC 的, 或 Ω 在 $\bar{t}$ 是 SNC 的. 则对任意的 $\mu\geqslant\ell_\varphi\cdot\ell$, 存在 $z^*\in Z^*$ 满足 $\|z^*\|\leqslant\mu$ 使得

$$0\in\partial\varphi(\bar{t})+D^*_NF(\bar{t},0)(z^*)+N(\bar{t};\Omega).$$

特别地, 如果 F 以

$$F(t):=g(t)+\Theta\quad (\text{其中}\ g:T\to Z,\ \Theta\subset Z)$$

形式给出, 则存在 $z^* \in -N(-g(\bar{t});\Theta)$ 满足 $\|z^*\| \leqslant \mu$, 使得

$$0 \in \partial\varphi(\bar{t}) + D_N^* g(\bar{t})(z^*) + N(\bar{t};\Omega), \tag{5.70}$$

其中假设 g 在 $\bar{t}$ 附近是连续的, Θ 在 $-g(\bar{t})$ 附近是局部闭的, 规范条件

$$D_M^* g(\bar{t})(0) \cap (-N(\bar{t};\Omega)) = \{0\} \tag{5.71}$$

成立, g 在 $\bar{t}$ 是 PSNC 的或 Ω 在该点是 SNC 的.

证明 从必要最优性条件的角度, 引理 5.47 中的惩罚最优化问题可等价地写为

$$\min \varphi(t) + \mu\|z\| \quad \text{s.t.} \quad z \in F(t), \quad t \in \Omega,$$

它是一般 MPEC(5.52) 的一种特殊形式. 在 Lipschitz 费用函数的情形下把定理 5.34(i) 的结果应用于此问题, 然后对 $\varphi(t) + \mu\|z\|$ 应用定理 2.33(c) 中的次微分和法则, 就证得定理的第一部分.

现在令 $F(t) := g(t) + \Theta$, 接着应用定理的一般陈述于这个特殊的映射, 它是 g 与 $\Theta(t) := \Theta$ $(\forall t \in T)$ 的和. 易见第二个映射在任意 $(\bar{t},\bar{z}) \in T \times \Theta$ 处是 PSNC 的, 并且它的两个上导数 $D^* = D_N^*, D_M^*$ 皆可如下计算:

$$D^*\Theta(\bar{t},\bar{z})(z^*) = \begin{cases} 0, & \text{若} -z^* \in N(\bar{z};\Theta), \\ \varnothing, & \text{其他}. \end{cases}$$

因此, 通过对和 $f + \Theta$ 的两个上导数 $D^* = D_N^*, D_M^*$ 应用定理 3.10 的上导数和法则, 有

$$D^*F(\bar{t},0)(z^*) \subset \begin{cases} D^*g(\bar{t})(z^*), & \text{若} -z^* \in N(-g(\bar{t});\Theta), \\ \varnothing, & \text{其他}. \end{cases}$$

把这个式子代入定理的一般规范和必要最优性条件中, 则分别得到关系 (5.71) 和 (5.70). 余下只需注意到根据定理 3.88, g 在 $\bar{t}$ 的 PSNC 性质蕴涵着 $F = g + \Theta$ 在 $(\bar{t},0)$ 的同样性质. △

注意到如果 g 在 $\bar{t}$ 附近是 Lipschitz 连续的, 则规范条件 (5.71) 成立, 并且它在该点是 PSNC 的. 还注意到基于精确惩罚的上面的方法, 不允许从 (5.52) 的上次微分最优化条件中推导出 (5.69) 的相应最优性条件, 这是由于所要求的和法则对和 $\varphi(\cdot) + \mu\|\cdot\|$ 的 Fréchet 上次微分来说一般不成立, 除非 φ 在最小值点是 Fréchet 可微的.

下面建立由参数变分系统控制的均衡约束的 MPEC 问题

$$\min \varphi(x,y) \quad \text{s.t.} \quad 0 \in f(x,y) + Q(x,y), \quad (x,y) \in \Omega \tag{5.72}$$

的必要最优性条件. 这里 $f\colon X\times Y\to Z$, $Q\colon X\times Y\rightrightarrows Z$, $\Omega\subset X\times Y$. 注意到问题 (5.72) 比 (5.56) 更一般, 那里的几何约束不依赖于 y. 下面所得的结果基于简约 MPEC 问题 (5.72) 为由非参数广义方程控制的 (5.69) 中的 MPEC 问题, 然后利用定理 5.48 的和分析法则. 注意到即使在对 y 独立的几何约束的情形下, 这些结果一般来说也不同于在 5.2.2 小节中所得的结果.

至少有两种方法把 (5.72) 化归为 (5.69). 第一种方法是直接考虑

$$F(t)=F(x,y):=f(x,y)+Q(x,y),$$

然后应用定理 5.48 中的一般最优化条件. 第二种方法是简化 (5.72) 为定理 5.48 中的一种特殊形式, 其中

$$\begin{aligned}F(x,y)&:=g(x,y)+\Theta,\quad \Theta:=\operatorname{gph}Q,\\ g(x,y)&:=\big(-x,-y,f(x,y)\big).\end{aligned}\tag{5.73}$$

为简洁起见, 只研究第二种方法. 它导致关于 MPEC 问题 (5.72) 的下面的必要最优性条件.

定理 5.49(通过惩罚的 MPEC 的最优性条件) 设 $(\bar x,\bar y)$ 是问题 (5.72) 的局部最优解, 其中 $f\colon X\times Y\to Z$, $Q\colon X\times Y\rightrightarrows Z$ 是 Asplund 空间之间的映射. 假设 φ 在 $(\bar x,\bar y)$ 附近是 Lipschitz 连续的, 且具有模 ℓ_φ, f 在该点附近是连续的, 假设集合 Ω 和 $\operatorname{gph}Q$ 分别在 $(\bar x,\bar y)$ 和 $(\bar x,\bar y,\bar z)$ 附近是局部闭的, 其中 $\bar z:=-f(\bar x,\bar y)$. 还假设由

$$G(u,v,w):=\big\{(x,y)\in\Omega\;\big|\;\big(u+x,v+y,w-f(x,y)\big)\in\operatorname{gph}Q\big\}$$

给出的映射 $G\colon X\times Y\times Z\rightrightarrows X\times Y$ 在 $(0,0,0,\bar x,\bar y)$ 是平静的, 且具有模 ℓ, 规范条件

$$D_M^*f(\bar x,\bar y)(0)\cap\Big(-N\big((\bar x,\bar y);\Omega\big)\Big)=\{0\}$$

成立, 并且 f 在 $(\bar x,\bar y)$ 是 PSNC 的或 Ω 在该点是 SNC 的. 则存在

$(x^*,y^*,z^*)\in X^*\times Y^*\times Z^*$ 满足 $\|(x^*,y^*,z^*)\|\leqslant\ell_\varphi\cdot\ell$ 和 $(x^*,y^*)\in D_N^*Q(\bar x,\bar y,\bar z)(z^*)$

使得

$$(-x^*,-y^*)\in\partial\varphi(\bar x,\bar y)+D_N^*f(\bar x,\bar y)(z^*)+N\big((\bar x,\bar y);\Omega\big),$$

此式蕴涵着

$$0\in\partial\varphi(\bar x,\bar y)+D_N^*f(\bar x,\bar y)(z^*)+D_N^*Q(\bar x,\bar y,\bar z)(z^*)+N\big((\bar x,\bar y);\Omega\big).$$

证明 应用定理 5.48 在 (5.73) 的数据下的特殊情形. 由于

$$g(x,y)=(-x,-y,0)+\big(0,0,f(x,y)\big),$$

由定理 1.70 易知, g 在 $(\bar{x},\bar{y})$ 是 PSNC 的当且仅当 f 在该点是 PSNC 的. 于是对两种上导数 $D^*=D^*_N, D^*_M$ 应用定理 1.62(ii) 中的和法则, 有

$$D^*g(\bar{x},\bar{y})(x^*,y^*,z^*)=(-x^*,-y^*)+D^*f(\bar{x},\bar{y})(z^*).$$

因此直接由定理 5.48 中的 (5.71) 和 (5.70) 得定理的规范条件和必要最优性条件. △

为进一步应用定理 5.49, 需给出有效条件确保此定理中映射 G 的平静性性质成立. 如所周知, 如果 G 在参考点附近是类 Lipschitz 的, 则它在该点是平静的. 由于 G 以约束系统的形式给出, 故平静性性质的充分条件由 4.3.2 小节中的结果可得. 下面利用这些结果; 简单起见, 只考虑 (5.72) 的均衡约束中的基 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的情形, 此时 f 在 $(\bar{x},\bar{y})$ 自动是 PSNC 的, 并且定理 5.49 中的规范条件满足; 因此 G 的类 Lipschitz 性质蕴涵着上面定理中 MPEC 问题的必要最优性条件.

推论 5.50(具有严格 Lipschitz 基的均衡约束) 在定理 5.49 的一般框架下, 假设 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, Q 在 $(\bar{x},\bar{y},\bar{z})$ 是 SNC 的, 关系

$$(x^*,y^*)\in\Big[\partial\langle z^*,f\rangle(\bar{x},\bar{y})+N\big((\bar{x},\bar{y});\Omega\big)\Big]\cap\big(-D^*_NQ(\bar{x},\bar{y},\bar{z})(z^*)\big) \tag{5.74}$$

只当 $(x^*,y^*,z^*)=(0,0,0)$ 时成立. 则存在 $z^*\in Z^*$ 使得必要最优性条件

$$0\in\partial\varphi(\bar{x},\bar{y})+\partial\langle z^*,f\rangle(\bar{x},\bar{y})+D^*_NQ(\bar{x},\bar{y},\bar{z})(z^*)+N\big((\bar{x},\bar{y});\Omega\big)$$

满足.

证明 设 $h\colon X\times Y\times Z\times X\times Y\to X\times Y\times Z$ 定义为

$$h(u,v,w,x,y):=\big(u+x,v+y,w-f(x,y)\big).$$

则定理 5.49 中的映射 G 可表示为约束系统

$$G(u,v,w)=\big\{(x,y)\in X\times Y\;\big|\;h(u,v,w,x,y)\in\operatorname{gph}Q,(u,v,w,x,y)\in X\times Y\times Z\times\Omega\big\}.$$

为保证 G 在 $(0,0,0,\bar{x},\bar{y})$ 附近的类 Lipschitz 性质, 应用推论 4.41 中的结果. 从 G 的结构易见, h 在 $(0,0,0,\bar{x},\bar{y})$ 是严格 Lipschitz 的, 当且仅当 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的并且集合 $\{0\}\times N((\bar{x},\bar{y});\Omega)$ 在 $(0,0,0,\bar{x},\bar{y})$ 相对于乘积空间 $X\times Y\times Z\times X\times Y$ 中的前三个分量是 PSNC 的. 于是对映射 G 应用推论 4.41 中的规范条件 (4.44), 得 $(u^*,v^*,w^*,x^*,y^*,z^*)=(0,0,0,0,0,0)$ 是包含系统

$$\begin{cases}(u^*,v^*,w^*,0,0)\in\partial\langle(x^*,y^*,z^*),h\rangle(0,0,0,\bar{x},\bar{y})+\{0\}\times N\big((\bar{x},\bar{y});\Omega\big),\\[2mm](x^*,y^*,z^*)\in N\big((\bar{x},\bar{y},\bar{z});\operatorname{gph}Q\big)\end{cases}$$

的唯一解, 这相当于要求由上面的关系有 $(x^*,y^*,z^*)=(0,0,0)$. 利用基本的次微分和法则, 有

$$\partial\big\langle(x^*,y^*,z^*),h\big\rangle(0,0,0,\bar{x},\bar{y})=\big(x^*,y^*,z^*,(x^*,y^*)+\partial\langle -z^*,f\rangle(\bar{x},\bar{y})\big),$$

因此上面的规范条件等于说系统（5.74）只有平凡解 $(x^*,y^*,z^*)=(0,0,0)$. 这就完成了推论的证明. △

与前一小节类似, 如果控制 (5.72) 中均衡约束的广义变分不等式具有复合势函数和复合域, 即当

$$Q(x,y)=\partial(\psi\circ g)(x,y)\ \ 和\ \ Q(x,y)=(\partial\psi\circ g)(x,y)$$

时, 基于分析法则可得到定理 5.49 和推论 5.50 的进一步推论. 用这种方法所得的结果由增广实值函数 ψ 的二阶次微分来表达; 这里把它留给读者. 对有限维空间中的一类特殊的 MPEC, 下面给出定理 5.49 的另一个推论, 其中定理 5.49 中的映射 G 可能不是类 Lipschitz 的, 但满足更弱的平静性性质.

考虑下面的具有多面体型的变分和非变分约束下的 MPEC 问题:

$$\begin{aligned}&\min\quad \varphi(x,y)\\&\text{s.t.}\quad 0\in A_1x+B_1y+c_1+Q(A_2x+B_2y+c_2),Lx+My+e\leqslant 0,\end{aligned}\tag{5.75}$$

其中 $x\in\mathbb{R}^n$, $y\in\mathbb{R}^m$, $e\in\mathbb{R}^p$, $Q\colon\mathbb{R}^s\rightrightarrows\mathbb{R}^m$, $C\in\mathbb{R}^{p\times n}$, $D\in\mathbb{R}^{p\times m}$, A_i, B_i, c_i $(i=1,2)$ 是适当维数的矩阵和向量.

推论 5.51(具有多面体约束的 MPEC 的最优性条件) 设 $(\bar{x},\bar{y})$ 是 (5.75) 的局部最优解. 假设 φ 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, Q 是分片多面体函数, 即它的图是有限个多面体集合的并. 则存在向量 $(x^*,y^*)\in\partial\varphi(\bar{x},\bar{y}),(u^*,v^*)\in\mathbb{R}^s\times\mathbb{R}^m,z^*:=(\lambda_1,\cdots,\lambda_p)\in\mathbb{R}^p$ 满足关系

$$0=x^*+A_2^*u^*+A_1^*v^*+L^*z^*,\quad 0=y^*+B_2^*u^*+B_1^*v^*+M^*z^*,$$

$$u^*\in D^*Q(A_2\bar{x}+B_2\bar{y}+c_2,-A_1\bar{x}-B_1\bar{y}-c_1)(v^*),$$

$$\lambda_j\geqslant 0,\quad \lambda_j\big((L\bar{x})_j+(M\bar{y})_j+e_j\big)=0,\quad j=1,\cdots,p.$$

证明 如上所述, 分片多面体集在它的有效域内任一点处是平静的. 由于两个集合 $\mathrm{gph}\,Q$ 和 $\varOmega:=\{(x,y)\in\mathbb{R}^{n+m}\mid Lx+My+e\leqslant 0\}$ 都是分片多面体集, 则映射 $G(u,v,w):=\{(x,y)\in\varOmega\mid(u+A_2x+B_2v+c_2,w-A_1x-B_1y-c_1)\in\mathrm{gph}\,Q\}$ 在 $(0,0,0)$ 是平静的, 因此定理 5.49 中的所有假设均成立. 而且考虑到作为 (5.72) 的

一种特殊形式的 (5.75) 中初始数据的特殊结构, 由定理 5.49 的必要最优性条件可直接推出推论的必要最优性条件成立. △

为说明所得的 MPEC 的必要最优性条件, 考虑下面的例子:

$$\min \frac{1}{2}x - y \quad \text{s.t.} \quad 0 \in y - x + \partial|y|, \quad x \in [-2, 0],$$

其中的均衡约束由所谓的第二种类型的一维变分不等式控制, 即由一凸连续函数的次微分定义.

易知 $(\bar{x}, \bar{y}) = (-1, 0)$ 是此问题的唯一全局解, 这是当 $Q(y) = \partial|y|$ 时 (5.75) 的一种特殊情形. 由于这个映射 Q 显然是分片多面体的, 推论 5.51 的所有假设都满足. 为验证这个推论的必要最优性条件, 需计算上导数 $D^*Q(0,1)$, 即在参考点处 $\partial|y|$ 的图的基本法锥. 从几何的角度利用定理 1.6 这个基本法锥可很容易地求得, 即有

$$N\big((0,-1); \operatorname{gph} \partial|\cdot|\big) = \big\{(u,v) \in \mathbb{R}^2 \,\big|\, uv = 0 \ \text{或} \ u > 0 \text{ 且 } v < 0\big\}.$$

于是推论 5.51 的必要最优性条件简化为

$$\left(\frac{1}{2}, -\frac{1}{2}\right) \in N\big((0,-1); \operatorname{gph} \partial|\cdot|\big),$$

这显然是满足的. △

注 5.52(MPEC 最优性条件的实现)　在上面特殊类型 MPEC 的最优性条件的应用中, 最具挑战性的任务是计算域多值函数 Q 的上导数 (或得到上导数的有效上估计). 在这种情况下, 当 Q 以次微分 $Q(\cdot) = \partial\psi(\cdot)$ 形式和以 5.22 小节中的复合次微分形式给出时, 就简化为计算或估计相应势函数的二阶次微分. 关于这样计算的一些例子和讨论已在 1.3.5 小节和 4.4.2 小节中给出; 特别地, 关于在力学中的应用, 参见例 4.67. 如上所述, 在最优化和各种应用中起着重要作用的一般非光滑函数类的二阶次微分已在 Dontchev 与 Rockafellar 的文章 [364] 和 Mordukhovich 与 Outrata 的文章 [939] 中计算. 这个方向的许多特殊计算和应用可在 Kočvara, Kružik 与 Outrata[689], Kočvara 与 Outrata[691, 690], Lucet 与 Ye[816], Mordukhovich, Outrata 与 Ĉervinka[940], Outrata[1024, 1025, 1027, 1030], Ye[1338, 1339], Ye 与 Ye[1343], Zhang[1360] 等文献及其参考文献中找到.

特别地, 对由

$$f(x,y) \geqslant 0, \quad y - g(x,y) \geqslant 0, \quad \big\langle f(x,y), y - g(x,y)\big\rangle = 0$$

给出的具有隐含互补约束的 MPEC 的完全计算已由 Outrata[1027] 完成, 其中 f 和 g 是从 $\mathbb{R}^n \times \mathbb{R}^m$ 到 $\mathbb{R}^m$ 的光滑单值映射. 这样的问题对各种工程、经济和力学的

应用来说非常重要. 它们相当于当 $g=0$ 时的标准非线性互补问题. 易见隐含互补约束可等价地写为均衡约束:

$$0 \in f(x,y) + (\partial\varphi \circ h)(x,y),$$

其中 $\varphi(\cdot) := \delta(\cdot;\mathbb{R}^m_+)$, $h(x,y) := y - g(x,y)$. 对 MPEC 的计算主要部分在于计算 $N(\cdot;\mathbb{R}^m_+)$ 的图的基本法锥, 这已由 Outrata 在文献 [1024] 中完成. 基于上面建立的非光滑分析法则, 可以把这些结果推广到具有非可微映射 f 和 g 的非光滑互补问题中.

5.3 多目标最优化

这一节致力于多目标约束最优化问题, 其中目标/费用函数可能不是实值的, 即最优化受控于相对更一般的偏好关系. 这样的问题, 可能首次出现在经济学的模型中 (例如, 参见第 8 章), 对于应用来说无疑非常重要. 在数学上它们也是令人感兴趣的问题, 与单目标极小化/极大化问题相比, 在大多数情况下有重大差别, 需要特殊的考虑.

下面研究无限维空间中具有各种约束的一般类多目标/向量最优化问题. 涉及的最优性 (有效性、均衡) 的概念由偏好关系给出, 包括在理论和应用中公认的标准关系, 同时在各种方向上进行了推广.

首先, 考虑这样一类多目标问题, 其中 Banach 空间之间的费用映射 $f\colon X \to Z$ 的最优性的概念用由给定子集 $\Theta \subset Z$ 定义的广义序关系描述, 该子集一般来说是非锥的和非凸的, 具有空的内部. 这样的 (f,Θ)- 最优性的概念实际上源于集合系统的局部极点的概念 (参见 2.1 节), 并且推广了经典的 Pareto/弱 Pareto 最优性及其扩展. 为得到具有各种约束的这种类型的多目标问题的必要最优性条件, 本节利用 2.2 节的极点原理以及本书发展的广义微分和 SNC 分析法则, 可得到这样的多目标及相关极小极大问题的由基本法向量和次梯度描述的广泛的结果. 值得注意的是, 这里的方法不依赖于任何标量化技术和在研究多目标最优化问题时通常使用的结果.

除上面类型的多目标问题以外, 本节还考虑一些类型的约束问题, 其中的最优性概念通常由满足某种传递性和局部饱足要求的抽象非自反偏好关系描述. 这样的偏好关系可能远远超出了广义 Pareto/弱 Pareto 最优性的范畴, 对某些重要应用来说是有益的. 为处理后面这种类型的多目标问题, 下面建立了推广的极点原理, 不仅应用于集合系统, 而且可应用于集值映射系统. 粗略地说, 通常的与推广的极点原理的主要区别在于, 在极点系统中推广的极点原理允许考虑集合的局部形变, 而不只是它们的 (线性) 平移. 在合理的假设下, 本节用这种方法建立了具有一般

非自反偏好关系的约束多目标问题的新的必要最优性条件. 另外还讨论所得结果的一些具体化及与以前结果的关系.

5.3.1　多目标问题的最优解

本节由抽象的最优性概念的定义开始, 它包括多目标问题的最优解的传统概念, 源于定义 2.1 中的集合极性的概念.

定义 5.53(广义序最优性)　给定 Banach 空间之间的单值映射 $f\colon X \to Z$, 集合 $0 \in \Theta \subset Z$, 称点 $\bar{x} \in X$ 是局部 (f,Θ)- 最优的, 如果存在 $\bar{x}$ 的邻域 U, 序列 $\{z_k\} \subset Z$ 满足当 $k \to \infty$ 时 $\|z_k\| \to 0$, 使得

$$f(x) - f(\bar{x}) \notin \Theta - z_k, \quad \forall x \in U, \quad k \in \mathbb{N}. \tag{5.76}$$

定义 5.53 中的集合 Θ 可以看成是由 $z_1 - z_2 \in \Theta$ 定义的 $z_1, z_2 \in Z$ 之间的推广序/偏好关系的一个生成集. 在 $Z = \mathbb{R}, \Theta = \mathbb{R}_-$ 的标量情形, 上面的最优性概念显然简化为标准的局部最优性.

值得注意的是, 通常不假设 Θ 是凸的或它的内部是非空的. 如果 Θ 是 Z 的凸子锥且 $\mathrm{ri}\Theta \neq \varnothing$, 则上面的最优性概念包括传统的最优性概念 (有时称为 Slater 最优性), 它要求不存在 $x \in U$ 满足 $f(x) - f(\bar{x}) \in \mathrm{ri}\Theta$. 而这种最优性又推广了弱 Pareto 最优性/有效性的概念, 它对应于 $f(x) - f(\bar{x}) \in \mathrm{int}\,\Theta$ 的情形. 事实上, 为将其简化为定义 5.53 中的概念, 可对某个 $z_0 \in \mathrm{ri}\,\Theta$, 在 (5.76) 中取 $z_k := -z_0/k \ \ (k \in \mathbb{N})$. Pareto 最优性的标准概念可用这些术语描述为不存在 $x \in U$, 使得 $f(x) - f(\bar{x}) \in \Theta$ 和 $f(\bar{x}) - f(x) \notin \Theta$. 当然, 当 $\Theta = \mathbb{R}^m_-$ 时, Pareto 型概念可用效用函数的经典术语描述.

另一方面, 把下面的紧集上的极小极大问题描述为多目标最优化问题对进一步研究来说是方便的.

例 5.54(极小极大描述为多目标最优化)　设 $\bar{x}$ 是极小极大问题:

$$\min \varphi(x) := \max\big\{\langle z^*, f(x)\rangle \,\big|\, z^* \in \Lambda\big\}, \quad x \in X$$

的局部最优解, 其中 $f\colon X \to Z, \Lambda \subset Z^*$ 是 Z^* 的弱 * 列紧子集, 使得存在 $z_0 \in Z$ 满足对任意的 $z^* \in \Lambda$ 有 $\langle z^*, z_0\rangle > 0$. 简单起见, 假设 $\varphi(\bar{x}) = 0$. 则 $\bar{x}$ 在定义 5.53 意义下是局部 (f,Θ) 最优的, 其中

$$\Theta := \big\{z \in Z \,\big|\, \langle z^*, z\rangle \leqslant 0, \forall z^* \in \Lambda\big\}.$$

证明　取 z_0 如上, 当序列 $z_k := z_0/k \ \ (k \in \mathbb{N})$ 时易验证 (5.76) 成立.　△

下一小节将表明在一般的约束下 (f,Θ)- 最优性可完全根据 2.2 节中的极点原理, 引理 5.23 中所得的广义上图的基本法向量的有效表示以及无限维空间中的 SNC 分析法则来处理.

然而, 对包括控制和对策理论框架中的多目标问题等的应用中, 其最优性的恰当概念需要的是如定义 5.53 中极点系统的集合的非线性变换而不是它们的线性平移. 这可通过考虑 Z 上满足某种要求的一般偏好关系来描述, 而这些要求允许使用变分分析中的合适技术.

给定子集 $Q \in Z \times Z$, 称 z_1 优先于 z_2, 并记为 $z_1 \prec z_2$, 如果 $(z_1, z_2) \in Q$. 序 $\prec$ 是非自反的, 如果相应的集合 Q 不含对角线 (z, z). 本书考虑满足下面条件的非自反序关系.

定义 5.55(闭偏好关系) 设

$$\mathcal{L}(z) := \{u \in Z \mid u \prec z\}$$

是在 $z \in Z$ 处相对于给定序 $\prec$ 的水平集. 称 $\prec$ 在 $\bar{z}$ 附近是局部饱足的, 如果对 $\bar{z}$ 的某邻域中的任意 z, 有 $z \in \operatorname{cl} \mathcal{L}(z)$. 而且, 称 $\prec$ 在 Z 上是几乎可传递的, 如果对任意的 $u \prec z, v \in \operatorname{cl} \mathcal{L}(u)$, 有 $v \prec z$. 序关系 $\prec$ 称为在 $\bar{z}$ 附近是闭的, 如果它同时是局部饱足的和几乎可传递的.

值得注意的是, 对任意合理的序局部饱足性质当然都是成立的, 但对在应用中非常重要的某些自然的序来说, 几乎传递性却可能被破坏, 尤其对与定义 5.53 中的 (f, Θ)- 最优性有关的那些序. 事实上, 可考虑所谓的"广义 Pareto"的情形, 其偏好关系由闭锥 $\Theta \subset Z$ 诱导, $z_1 \prec z_2$ 当且仅当 $z_1 - z_2 \in \Theta$ 且 $z_1 \neq z_2$. 当然这是定义 5.53 的特殊情形. 下面的命题完全描述了这个序是几乎可传递时对 Θ 的要求. 回顾一下概念, 称锥 Θ 是尖的, 如果 $\Theta \cap (-\Theta) = \{0\}$.

命题 5.56(几乎可传递的广义 Pareto 序) 如上定义的广义 Pareto 序 $\prec$ 是几乎可传递的, 当且仅当锥 $\Theta \subset Z$ 是凸的和尖的.

证明 首先证明, 如果上面的序 $\prec$ 是几乎可传递的, 则锥 Θ 是凸的. 任取 $z_1, z_2 \in \Theta \backslash \{0\}, \lambda \in (0,1)$ 和 $a \in Z$, 定义 $u := a + \lambda z_1, v := a - (1-\lambda) z_2$. 由于 $\lambda z_1 \neq 0$, 有 $u \prec a$ 和 $a \prec v$. 根据几乎传递性有 $u \prec v$, 这意味着 $\lambda z_1 + (1-\lambda) z_2 = u - v \in \Theta$, 即 Θ 是凸的.

为证在 $\prec$ 的传递性条件下 Θ 是尖的, 取 $z \in \Theta \cap (-\Theta)$, 令 $u := a + z, v := a - (-z)$. 如果 $z \neq 0$, 则几乎传递性蕴涵着 $u \prec v$, 从而 $0 = u - v \in \Theta \setminus \{0\}$, 矛盾, 因此 $z = 0$.

为证命题相反的陈述成立, 假设 Θ 是凸的和尖的, 取 $v \in \operatorname{cl} \mathcal{L}(u)$ 且 $u \prec z$. 则存在 $z_1, z_2 \in \Theta$, 满足 $v = u + z_1, z = u - z_2$ 且 $z_2 \neq 0$. 由 Θ 的凸性得 $(v-z)/2 = z_1/2 + z_2/2 \in \Theta$, 因此 $v \in \operatorname{cl} \mathcal{L}(z)$. 由对 $v = z$ 的假设得 $z_1 = -z_2 \neq 0$, 这与 Θ 的尖性相矛盾. 因此有 $v \prec z$, 并完成了命题的证明. △

由命题 5.56 的刻画可以看到, 对广义 Pareto 序的某些重要特殊情形 (从而也包括定义 5.53 的情形), 定义 5.55 中的几乎传递性条件可能不成立. 特别地, 下面

$\mathbb{R}^m$ 上的字典序描述的序关系就是这种情形.

例 5.57(字典序) 设 $\prec$ 是 $\mathbb{R}^m$ 上由字典序定义的一个序关系, $m \geqslant 3$, 即 $u \prec v$, 如果存在一整数 $j \in \{0, \cdots, m-1\}$, 使得对向量 $u, v \in \mathbb{R}^m$ 的相应分量, $u_i = v_i$ 对 $i = 1, \cdots, j$ 成立, 且 $u_{j+1} < v_{j+1}$. 则这个序关系在 $\mathbb{R}^m$ 上是局部饱足的, 但不是几乎可传递的.

证明 易验证字典序 $\prec$ 在 $\mathbb{R}^m$ 上是局部饱足的. 另一方面, 几乎传递性对该偏好关系是不成立的. 为说明这一点, 考虑 $\mathbb{R}^m$ 中的向量

$$z := (0, 0, 1, \cdots, 0), \quad u := (0, \cdots, 0), \quad v := (0, 1, 1, 0, \cdots, 0)$$

和序列 $v_k := (-1/k, 1, 1, 0, \cdots, 0) \to v(k \to \infty)$. 则 $u \prec z, v_k \prec u$, 但 $v \not\prec z$ 而 $v \in \operatorname{cl} \mathcal{L}(u)$. △

这一节余下的部分建立约束多目标问题的必要最优性条件, 其中 (向量值) 映射 $f \colon X \to Z$ 在 $\bar{x}$ 处的局部最优性的概念分别由定义 5.53 意义下 Z 上的一广义序 Θ 和定义 5.55 意义下 Z 上的闭序给出. 在这两种情形下所得的结果基于不太一样的技术, 一般来说是独立的.

5.3.2 广义序最优性

本小节研究约束多目标问题在定义 5.53 意义下的局部最优解的必要最优性条件. 定义 5.53 表明了应用集合系统的极点原理来得到这样广义序最优性必要条件的可能性, 实际上, 该定义是受集合系统的局部极点概念的启发而来的. 这里的主要目标是得到以点基/确切形式表示的涉及参考最优解处的广义微分结构的必要条件. 主要精力放在规范型的必要最优性条件; 利用对偶空间方法, 相应的非规范最优性条件类似可得.

为在最少的假设下得到广义序最优性必要条件的一般结果, 需要关于 Asplund 空间的乘积空间中两个集合情形的定理 2.22 中确切极点原理的推广版本. 这个结果包含相对于指标集 $J \subset \{1, 2\}$ 乘积空间 $X_1 \times X_2$ 中的集合 (有可能是空集) 的 PSNC 和强 PSNC 性质; 参见定义 3.3. 注意到如果 $J = \varnothing$, 则 PSNC 和强 PSNC 性质自动成立; 当 $J = \{1, 2\}$ 时, 这两种性质都化归为集合的 SNC 性质. 本节的兴趣所在是下面引理中的中间情形, 即考虑到乘积结构, 这对本小节的主要结果是非常重要的.

引理 5.58(Asplund 空间的乘积空间中的确切极点原理) 设 $\bar{x} \in \Omega_1 \cap \Omega_2$ 是集合 $\Omega_1, \Omega_2 \subset X_1 \times X_2$ 的局部极点, 且 Ω_1, Ω_2 在 $\bar{x}$ 附近是局部闭的, 设 $J_1, J_2 \subset \{1, 2\}$ 且 $J_1 \cup J_2 = \{1, 2\}$. 假设 X_1 和 X_2 都是 Asplund 空间, 而且 Ω_1 在 $\bar{x}$ 相对于 J_1 是 PSNC 的, Ω_2 在 $\bar{x}$ 相对于 J_2 是强 PSNC 的. 则存在 $x^* \neq 0$ 满足

$$x^* \in N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big).$$

证明 应用定理 2.20 的近似极点原理于极点系统 $\{\Omega_1, \Omega_2, \bar{x}\}$, 取序列 $\varepsilon_k \downarrow 0\ (k \to \infty)$, 找到 $u_k \in \Omega_1, v_k \in \Omega_2, u_k^* \in \widehat{N}(u_k; \Omega_1)$ 和 $v_k^* \in \widehat{N}(v_k; \Omega_2)$, 使得

$$\|u_k - \bar{x}\| < \varepsilon_k, \quad \|v_k - \bar{x}\| < \varepsilon_k, \quad \|u_k^* + v_k^*\| < \varepsilon_k,$$

$$\frac{1}{2} - \varepsilon_k < \|u_k^*\| < \frac{1}{2} + \varepsilon_k, \quad \frac{1}{2} - \varepsilon_k < \|v_k^*\| < \frac{1}{2} + \varepsilon_k$$

对任意 $k \in \mathbb{N}$ 成立. 由于序列 $\{u_k^*\}$ 和 $\{v_k^*\}$ 在 Asplund 空间中的对偶空间中是有界的, 它们沿子序列弱 * 收敛于某 u^* 和 v^*(不失一般性, 设原序列都分别收敛). 则 $u^* \in N(\bar{x}; \Omega_1)$, $v^* \in N(\bar{x}; \Omega_2)$, 并且 $u^* + v^* = 0$.

余下需证 $x^* := u^* \neq 0$. 假设相反, 应用 Ω_2 在 $\bar{x}$ 相对于 J_2 的强 PSNC 性质, 对每个 $j \in J_2$ 有 $\|v_{jk}^*\| \to 0 (k \to \infty)$. 由近似极点原理的关系知, 对任意 $j \in J_2$ 有 $\|u_{jk}^*\| \to 0$. 由于 Ω_1 被假定为在 $\bar{x}$ 相对于 J_1 是 PSNC 的, $\forall j \in J_1$, 有 $\|u_{jk}^*\| \to 0$. 由于 $J_1 \cup J_2 = \{1, 2\}$, 得 $\|u_k^*\| \to 0\ \ (k \to \infty)$, 矛盾, 从而完成了引理的证明. △

基于上面的引理和广义微分分析法则, 下面将得到约束多目标问题的必要最优性条件, 这里的最优性按定义 5.53 来理解. 给定映射 $f\colon X \to Z$, 集合 $\Omega \subset X, \Theta \subset Z$, 考虑在 (5.37) 中定义的广义上图 $\mathcal{E}(f, \Omega, \Theta)$ 和 f 在 Ω 上的限制 $f_\Omega := f|_\Omega$. 回顾下面定理的最后所需要的在定义 4.8 中引入的强上导数正规性的概念, 这个性质的一些充分条件在命题 4.9 中给出.

定理 5.59(广义序最优性的必要条件) 设 $f\colon X \to Z$ 是 Asplund 空间之间的映射, 集合 $\Omega \subset X, \Theta \subset Z$ 满足 $\bar{x} \in \Omega, 0 \in \Theta$. 假设点 $\bar{x}$ 相对于 Ω(即服从约束 $x \in \Omega$) 是局部 (f, Θ)- 最优的. 则下面的断言成立:

(i) 假设集合

$$\mathcal{E}(f, \Omega, \Theta) := \big\{(x, z) \in X \times Z \,\big|\, f(x) - z \in \Theta, x \in \Omega\big\}$$

在 $(\bar{x}, \bar{z})$ 附近是局部闭的, 其中 $\bar{z} := f(\bar{x})$, 且 $\dim Z < \infty$. 则存在 $z^* \in Z^*$ 满足

$$(0, -z^*) \in N\big((\bar{x}, \bar{z}); \mathcal{E}(f, \Omega, \Theta)\big), \quad z^* \neq 0, \tag{5.77}$$

这总蕴涵着 $z^* \in N(0; \Theta)$. 如果 f 在 $\bar{x}$ 附近相对于 Ω 是连续的, Ω 和 Θ 分别在 $\bar{x}$ 和 0 附近是局部闭的, 则 (5.77) 也蕴涵着 $0 \in D_N^* f_\Omega(\bar{x})(z^*)$. 进一步, 如果 f 在 $\bar{x}$ 点是 w^*- 严格 Lipschitz 的, 则 (5.77) 等价于

$$0 \in \partial\langle z^*, f_\Omega\rangle(\bar{x}), \quad z^* \in N(0; \Theta) \setminus \{0\}. \tag{5.78}$$

(ii) 设 f 在 $\bar{x}$ 附近相对于 Ω 是连续的, Ω 和 Θ 分别在 $\bar{x}$ 和 0 附近是闭的, 设下述条件之一成立:

(a)Θ 在 0 处是 SNC 的,

(b)f_Ω^{-1} 在 $(\bar{z},\bar{x})$ 是 PSNC 的.

则存在 $z^*\in Z^*$ 满足

$$0\neq z^*\in N(0;\Theta)\cap\ker D_N^*f_\Omega(\bar{x}),\tag{5.79}$$

如果 f 在 $\bar{x}$ 是 w^*- 严格 Lipschitz 的, f_Ω 在该点是强上导数正规的, 那么 (5.79) 等价于 (5.78) 和 (5.77).

证明 简单起见, 假设 $\bar{z}=f(\bar{x})=0$. 则根据定义 5.53, 点 $(\bar{x},0)\in X\times Z$ 是集合系统 $\{\Omega_1,\Omega_2\}$ 的局部极点, 这里

$$\Omega_1:=\mathcal{E}(f,\Omega,\Theta),\quad \Omega_2:=\operatorname{cl}U\times\{0\},$$

U 是 (5.76) 中局部最优点的一个邻域, 其中 $x\in\Omega$.

首先考虑断言 (i) 的情形, 其中集合 Ω_1 在 $(\bar{x},0)$ 附近是局部闭的, 即使 f 在 $\bar{x}$ 附近可能不是连续的; 比照定理 5.24. 由于 U 是 $\bar{x}$ 的邻域, Z 是有限维的, 则集合 Ω_2 在 $(\bar{x},0)$ 是 SNC 的. 此时可应用定理 2.22 中确切极点原理的通常版本, 立得 (5.77), 从而 $z^*\in N(0;\Theta)$. (i) 中的另一结论在所给假设下由引理 5.23 可得.

接下来考虑断言 (ii) 中一般 Asplund 空间情形, 其中 f 的连续性假设, 以及 Ω 和 Θ 的闭性假设直接蕴涵着 $\mathcal{E}(f,\Omega,\Theta)$ 在 $(\bar{x},0)$ 附近的局部闭性. 因此如果存在指标集 $J_1,J_2\subset\{1,2\}$ 且 $J_1\cup J_2=\{1,2\}$ 满足 Ω_1 在 $(\bar{x},0)$ 相对于 J_1 是 PSNC 的, 而 Ω_2 在 $(\bar{x},0)$ 相对于 J_2 是强 PSNC 的, 可利用引理 5.58 中确切极点原理的乘积空间版本.

取 $J_1=\{2\},J_2=\{1\}$, 即在引理 5.58 中, 取 $X_1=Z,X_2=X$. 易见, 由于 U 是 $\bar{x}$ 的邻域, Ω_2 在 $(\bar{x},0)$ 相对于 X 是强 PSNC 的; 注意到 Ω_2 在 $(\bar{x},0)$ 从不是 SNC 的, 除非 Z 是有限维的. 余下的只需证明集合 $\Omega_1=\mathcal{E}(f,\Omega,\Theta)$ 在 $(\bar{x},0)$ 相对于 Z 是 PSNC 的. 由于 Ω_1 可表为逆像:

$$\mathcal{E}(f,\Omega,\Theta)=g^{-1}(\Theta),\quad 其中\quad g(x,z):=f_\Omega(x)-z,$$

利用定理 3.84 可保证这个集合的强 SNC 性质. 然而, 用这种方法得到所需要的 PSNC 性质则会要求过多的条件. 下面根据映射 g 的特殊结构来建立满足 PSNC 性质的更多精细充分条件. 这里需证明, 任给序列 $(x_k,z_k)\to(\bar{x},0), x_k\in\Omega, g(x_k,z_k)\in\Theta$ 和 $(x_k^*,z_k^*)\in\widehat{N}((x_k,z_k);\mathcal{E}(f,\Omega,\Theta))$, 有

$$\left[\|x_k^*\|\to 0, z_k^*\xrightarrow{w^*}0\right]\Rightarrow\|z_k^*\|\to 0.$$

考虑由 $\Lambda_1:=\operatorname{gph}g,\Lambda_2:=X\times Z\times\Theta$ 定义的局部闭集 $\Lambda_1,\Lambda_2\subset X\times Z\times Z$, 并且注意到

$$(x_k^*,z_k^*,0)\in\widehat{N}\big((x_k,z_k,v_k);\Lambda_1\cap\Lambda_2\big),\quad\forall k\in\mathbb{N},\tag{5.80}$$

这里 $v_k := g(x_k, z_k)$. 利用关于三个空间的乘积 $X \times Z \times Z = X_1 \times X_2 \times X_3$ 中交集的 PSNC 性质的定理 3.79, 下面证明集合 $\Lambda_1 \cap \Lambda_2$ 在 $(\bar{x}, 0, 0)$ 相对于 $Z = X_2$ 是 PSNC 的.

首先考虑当 Θ 被假定为在 0 点是 SNC 时 (ii) 中的情形 (a). 此时按照定理 3.79 的记号取 $J_1 = \{2\}, J_2 = \{1, 2, 3\}$, 并注意到 Λ_2 在 $(\bar{x}, 0, 0)$ 是 SNC 的, 从而它在 $(\bar{x}, 0, 0)$ 相对于 $J_2 \setminus J_1 = \{1, 3\}$ 的强 PSNC 性质自动成立. 下面验证 Λ_1 在 $(\bar{x}, 0, 0)$ 相对于 $J_1 = \{2\}$ 是 PSNC 的, 即对任意序列 $(x_k, z_k, v_k) \to (\bar{x}, 0, 0), (x_k^*, z_k^*, v_k^*) \in \widehat{N}((x_k, z_k, v_k); \operatorname{gph} g)$, 当 $k \to \infty$ 时, 有

$$\left[\|(x_k^*, v_k^*)\| \to 0, z_k^* \xrightarrow{w^*} 0\right] \Rightarrow \|z_k^*\| \to 0.$$

事实上, 由于关系 $(x_k^*, z_k^*, v_k^*) \in \widehat{N}((x_k, z_k, v_k); \operatorname{gph} g)$ 可以重写为 $(x_k^*, z_k^*) \in \widehat{D}^* g(x_k, z_k)(-v_k^*)$, 故根据 g 的结构和定理 1.62(i) 有

$$x_k^* \in \widehat{D}^* f_\Omega(x_k)(-v_k^*), \quad z_k^* = v_k^*. \tag{5.81}$$

由 (5.81) 得 $\|z_k^*\| = \|v_k^*\| \to 0\,(k \to \infty)$, 这就证明了 Λ_1 在 $(\bar{x}, 0, 0)$ 相对于 $J_1 = \{2\}$ 的 PSNC 性质.

为在情形 (a) 下利用定理 3.79, 余下需证集合系统 $\{\Lambda_1, \Lambda_2\} \subset X_1 \times X_2 \times X_3$ 在 $(\bar{x}, 0, 0)$ 相对于 $(J_1 \backslash J_2) \cup (J_2 \backslash J_1) = \{1, 3\}$ 定义 3.78 中的混合规范条件成立. 根据 Λ_1 和 Λ_2 的结构, 该混合规范条件简化为：对任意序列 $(x_k^*, z_k^*, v_k^*) \in \widehat{N}((x_k, z_k, v_k); \operatorname{gph} g), u_k^* \in \widehat{N}(u_k; \Theta)$, 若 $k \to \infty$ 时对某个 $u^* \in N(0; \Theta)$ 满足

$$(x_k, z_k, v_k, u_k) \to (\bar{x}, 0, 0, 0), \quad \|x_k^*\| \to 0, \quad z_k^* \xrightarrow{w^*} 0, \quad (u_k^*, v_k^*) \xrightarrow{w^*} (u^*, -u^*),$$

则有 $u^* = 0$. 由前面的讨论可得, 上面的序列对任意的 $k \in \mathbb{N}$ 关系 (5.81) 成立. 因此所论的混合规范条件对所考虑的集合 Λ_1 和 Λ_2 来说等价于

$$N(0; \Theta) \cap \ker \widetilde{D}_M^* f_\Omega(\bar{x}) = \{0\}. \tag{5.82}$$

如果 (5.82) 不成立, 由于混合上导数不大于基本上导数, 则立即得最优性条件 (5.79). 因此可假设 (5.82) 成立, 于是定理 3.79 保证交集 $\Lambda_1 \cap \Lambda_2$ 在 $(\bar{x}, 0, 0)$ 相对于 Z 的 PSNC 性质. 根据这个 PSNC 性质由 (5.80) 得 $\|z_k^*\| \to 0$, 从而集合 $\mathcal{E}(f, \Omega, \Theta)$ 在 $(\bar{x}, 0)$ 相对于 Z 是 PSNC 的. 这就可以对上面的集合系统 $\{\Omega_1, \Omega_2\}$ 应用引理 5.58, 并得到最优性条件 (5.77). 进一步, 根据引理 5.23, (5.77) 蕴涵 (5.79), 这也保证在情形 (a) 下定理的其他结论成立.

为完成定理的证明, 余下需考虑 (ii) 中的情形 (b). 与在情形 (a) 下的证明的唯一差别在于现在需证, 若 f_Ω^{-1} 在 $(0, \bar{x})$ 是 PSNC 的, 则交集 $\Lambda_1 \cap \Lambda_2$ 在 $(\bar{x}, 0, 0)$ 相对

于乘积空间 $X \times Z \times Z = X_1 \times X_2 \times X_3$ 中的 $Z = X_2$ 是 PSNC 的. 为此在指标集的不同安排下再次应用定理 3.79, 即按定理 3.79 中的记号取 $J_1 = \{2,3\}, J_2 = \{1,2\}$. 则 $J_2 \backslash J_1 = \{1\}$, 而且集合 Λ_2 在 $(\bar{x},0,0)$ 相对于 J_2 显然是强 PSNC 的. 下面验证在 (b) 的假设下 Λ_1 在 $(\bar{x},0,0)$ 相对于 J_1 是 PSNC 的. 即需证明, 对任意序列 $(x_k, z_k, v_k) \to (\bar{x},0,0), (x_k^*, z_k^*, v_k^*) \in \widehat{N}\big((x_k, z_k, v_k); \operatorname{gph} g\big)$, 有

$$\Big[\|x_k^*\| \to 0, (z_k^*, v_k^*) \xrightarrow{w^*} (0,0)\Big] \Rightarrow \|(z_k^*, v_k^*)\| \to 0 \quad (k \to \infty).$$

根据情形 (a) 中的讨论, 对满足

$$x_k^* \in \widehat{D}^* f_\Omega(x_k)(z_k^*) \text{ 且 } x_k \to \bar{x}, \quad \|x_k^*\| \to 0, \quad z_k^* \xrightarrow{w^*} 0$$

的任意序列 (x_k, x_k^*, z_k^*), 有 $\|z_k^*\| \to 0 (k \to \infty)$, 这显然等价于 (b) 假设下 f_Ω^{-1} 在 $(0,\bar{x})$ 的 PSNC 性质.

最后, 为在情形 (b) 下应用定理 3.79, 需要定义 3.78 中的混合规范条件对 $\{\Lambda_1, \Lambda_2\}$ 在 $(\bar{x},0,0)$ 相对于 $(J_1 \backslash J_2) \cup (J_2 \backslash J_1) = \{1,3\}$ 成立, 这恰好与情形 (a) 相同. 这就完成了定理的证明. △

考虑到 $f_\Omega(x) = f(x) + \Delta(x;\Omega)$, 其中 $\Delta(x;\Omega)$ 是集合 Ω 的指示映射, 并利用第 3 章中建立的上导数/次微分与 PSNC 和法则, 由定理 5.59 易得分别用 f 和 Ω 表达的相应 (通常有更多限制的) 条件. 通过利用交集

$$\mathcal{E}(f,\Omega,\Theta) = \big\{(x,z) \,\big|\, f(x) - z \in \Theta\big\} \cap (\Omega \times Z)$$

的交法则, 在 (5.77) 的框架下, 这样的条件也能得到. 这就可以 (再次基于广泛的分析法则) 有效地处理关于广义序最优性的一类约束多目标问题, 其中的约束类似 5.1 节研究的单目标极小化的各种约束形式. 用这种方法得到具有表述为 $x \in G^{-1}(\Lambda) \cap \Omega$ 的几何和算子约束的多目标问题的必要最优性条件, 这特别地包含等式和不等式型的泛函约束. 下面给出情形 (a) 下定理 5.59 的一个推论, 其中为简单起见, 只考虑由在 Asplund 空间之间的集值映射下集合的逆像给出的算子 (没有几何) 约束的多目标问题.

推论 5.60(具有算子约束的多目标问题) 设 $f\colon X \to Z$, $G\colon X \rightrightarrows Y$ 是 Asplund 空间之间的映射, 设 $\Theta \subset Z, \Lambda \subset Y$ 是非空子集. 假设 $\bar{x}$ 受约束于 $x \in G^{-1}(\Lambda)$ 是 (f,Θ)- 最优的, 其中 f 是连续的, G 是闭图的, Θ 和 Λ 在相应点附近是闭的. 还假设 Θ 在 0 点是 SNC 的, 集值映射 $S(\cdot) := G(\cdot) \cap \Lambda$ 在 $\bar{x}$ 附近是内半紧的. 如果关于 (f,G,Λ) 的下列条件之一成立:

(a) f 在 $\bar{x}$ 是 PSNC 的, 规范条件

$$N(\bar{y};\Lambda) \cap \ker \widetilde{D}_M^* G(\bar{x},\bar{y}) = \{0\}, \quad \forall \bar{y} \in S(\bar{x}),$$

$$\cup\Big[D_N^*G(\bar{x},\bar{y})(y^*)\big|\bar{y}\in S(\bar{x}),y^*\in N(\bar{y};\Lambda)\Big]\cap\big(-D_M^*f(\bar{x})(0)\big)=\{0\}$$

成立, 且或者 G^{-1} 在 $(\bar{y},\bar{x})$ 是 PSNC 的, 或者对所有 $\bar{y}\in S(\bar{x})$ 映射 Λ 在 $\bar{y}$ 是 SNC 的.

(b) (a) 中的第二个规范条件和

$$N(\bar{y},\Lambda)\cap\ker D_N^*G(\bar{x},\bar{y})=\{0\},\quad \forall\,\bar{y}\in S(\bar{x})$$

成立, 另外, 或者 G 在 $(\bar{x},\bar{y})$ 是 PSNC 的且 Λ 在 $\bar{y}$ 是 SNC 的, 或者 G 对所有 $\bar{y}\in S(\bar{x})$ 在 $(\bar{x},\bar{y})$ 是 SNC 的.

则存在 $z^*\in N(0;\Theta)\setminus\{0\}$ 满足

$$\cup\Big[D_N^*G(\bar{x},\bar{y})(y^*)\big|\bar{y}\in S(\bar{x}),y^*\in N(\bar{y};\Lambda)\Big]\cap\big(-D_N^*f(\bar{x})(z^*)\big)\neq\varnothing.$$

证明 应用情形 (a) 下的定理 5.59(ii), 找到 $z^*\in N(0;\Theta)$ 满足

$$0\in D_N^*\big[f+\Delta(\cdot;\Omega)\big](\bar{x})(z^*),$$

其中 $\Omega:=G^{-1}(\Lambda)$. 根据定理 3.10 中的上导数和法则, 上式蕴涵着

$$N(\bar{x};G^{-1}(\Lambda))\cap\big(-D_N^*f(\bar{x})(z^*)\big)\neq\varnothing$$

在

$$N\big(\bar{x};G^{-1}(\Lambda)\big)\cap\big(-D_M^*f(\bar{x})(z^*)\big)=\{0\}$$

和 f 在 $\bar{x}$ 是 PSNC 的或 $G^{-1}(\Lambda)$ 在该点是 SNC 的条件下成立. 如果 f 在 $\bar{x}$ 是 PSNC 的, 则推论由定理 3.8 可得, 该定理给出了在 (a) 中所给的假设下, 逆像 $G^{-1}(\Lambda)$ 的基本法向量的表示. 否则, 需利用定理 3.84 保证 $G^{-1}(\Lambda)$ 在 $\bar{x}$ 的 SNC 性质, 在 (b) 中所给的假设下就得到推论的结论. △

值得注意的是, 如果 f 在 $\bar{x}$ 附近是 Lipschitz 连续的, 对所有 $\bar{y}\in S(\bar{x})$, 映射 G 在 $(\bar{x},\bar{y})$ 附近是度量正则的, 则关于 f 和 G^{-1} 的 PSNC 条件和推论 5.60(a) 中的两个规范条件都自动成立. 还注意到推论 5.60 的结果在最合适的约束规范下给出了多目标问题的规范型必要最优性条件. 它们蕴涵着没有约束规范的情况下非规范形式的必要条件; 请比照 5.1.2 小节关于极小化增广实值函数的问题 (即单目标的问题).

类似 5.1.3 小节, 根据定理 5.59 和推论 5.60 可得到由等式和不等式给出的泛函约束的多目标问题的相应必要最优性条件. 然而, 5.1.3 小节的一些结果 (和 5.1 节前面的材料) 从实质上利用了单目标极小化问题的一些特殊性质. 特别地, 它包含极小化增广实值函数的上次微分条件. 然而, 对多目标问题, 在关于所论 Asplund

空间 X 的额外假设下, 只可得到涉及不等式约束 (而不是目标函数) 的上次微分型的必要最优性条件. 简单起见, 下面只给出没有约束规范只有不等式约束的多目标问题的必要最优性条件, 从上次微分最优性条件的角度看, 只有这些约束才具有特殊意义.

定理 5.61(多目标问题的上次微分最优性条件) 给定 $f\colon X \to Z$, $\Theta \subset Z$ 在 0 附近是闭的, 假设 $\bar{x}$ 在不等式约束

$$\varphi_i(x) \leqslant 0, \quad i = 1, \cdots, m$$

下是 (f, Θ)- 最优的, 其中 $\varphi_i\colon X \to \overline{\mathbb{R}}$ 在 $\bar{x}$ 有限 $(i = 1, \cdots, m)$. 假设 Z 是 Asplund 空间, X 有 Lipschitz $\mathcal{C}^1$- 光滑阻尼函数 (当 X 有 Fréchet 可微重赋范时, 这个条件自动满足), f 在 $\bar{x}$ 附近是 Lipschitz 连续的, Θ 在原点是 SNC 的. 则对任意 Fréchet 上次梯度 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$, $i = 1, \cdots, m$, 存在 $z^* \in N(0; \Theta)$ 和 $(\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m$, 满足

$$\lambda_i \geqslant 0, \quad \lambda_i\varphi(\bar{x}) = 0, \quad i = 1, \cdots, m,$$

使得 $(z^*, \lambda_1, \cdots, \lambda_m) \neq 0$, 而且有

$$0 \in D_N^* f(\bar{x})(z^*) + \sum_{i=1}^{m} \lambda_i x_i^*.$$

进一步, 如果 f 在 $\bar{x}$ 是严格 Lipschitz 的, 则

$$0 \in \partial\langle z^*, f\rangle(\bar{x}) + \sum_{i=1}^{m} \lambda_i x_i^*, \quad (z^*, \lambda_1, \cdots, \lambda_m) \neq 0.$$

证明 给定 $x_i^* \in \widehat{\partial}\varphi_i(\bar{x})$, $i = 1, \cdots, m$, 当 $\mathcal{S} = \mathcal{LC}^1$(即空间中有 Lipschitz $\mathcal{C}^1$ 阻尼函数) 时, 应用定理 1.88(ii) 中的 Fréchet 次梯度 $-x_i^* \in \widehat{\partial}(-\varphi_i)(\bar{x})$ 的变分描述, 找到 $s_i\colon X \to \mathbb{R}$, $i = 1, \cdots, m$, $\bar{x}$ 的邻域 U, 使得

$$s_i(\bar{x}) = \varphi_i(\bar{x}), \quad s_i(x) \geqslant \varphi_i(x), \quad \forall x \in U, \quad i = 1, \cdots, m,$$

并且每个 s_i 在 U 上是连续可微的, 满足 $\nabla s_i(\bar{x}) = x_i^*$, $i = 1, \cdots, m$. 由 s_i 的构造可知, $\bar{x}$ 是 (f, Θ)- 最优解, 其中约束条件不仅包括原始约束 $\varphi_i(\bar{x}) \leqslant 0$, 还包括光滑不等式约束

$$s_i(x) \leqslant 0, \quad i = 1, \cdots, m.$$

当 $G(x) := (s_1(x), \cdots, s_m(x)), \Lambda := \mathbb{R}^m_-$ 时, 运用推论 5.60 中的必要最优性条件, 则有

$$D^*G(\bar{x})(\lambda_1, \cdots, \lambda_m) = \sum_{i=1}^{m} \lambda_i \nabla s_i(x) = \sum_{i=1}^{m} \lambda_i x_i^*,$$

$$N(\bar{y}; \Lambda) = \{(\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m \mid \lambda_i \geqslant 0, \lambda_i \varphi_i(\bar{x}) = 0, \quad i = 1, \cdots, m\},$$

这里 $\bar{y} := (\varphi_1(\bar{x}), \cdots, \varphi_m(\bar{x})) = (s_1(\bar{x}), \cdots, s_m(\bar{x}))$. 易见对所有考虑的问题, 除了关于 $D^*G(\bar{x})$ 的核的规范条件外, 推论 5.60(b) 中的所有其他假设均满足. 如果这个规范条件满足, 根据推论 5.60, 当 $z^* \neq 0$ 时, 得到定理的结论. 否则, 当 $(\lambda_1, \cdots, \lambda_m) \neq 0$ 时, 定理的结论显然成立, 这就完成了证明. △

根据指定的定序集 Θ, 由上面所得的一般结果可得特殊多目标问题的必要最优性条件. 注意到, 如果 Θ 是凸锥 (这是许多典型应用中的情形) 则定理 5.59, 定理 5.61 和推论 5.60 中的最优性条件 $z^* \in N(0; \Theta)$ 可表示为

$$\langle z^*, z \rangle \leqslant 0, \quad \forall z \in \Theta.$$

下面给出上面的结果在约束极小极大问题:

$$\min \varphi(x) := \sup \left\{ \langle z^*, f(x) \rangle \mid z^* \in \Lambda \right\} \quad \text{s.t.} \quad x \in \Omega \tag{5.83}$$

中的一个应用, 其中 $f\colon X \to Z$, $\Omega \subset X$, $\Lambda \subset Z^*$. 正如例 5.54 表明的那样, (5.83) 中的极小极大目标可简化为定理 5.59 中考虑的广义序最优性. 下面的结果给出了定理 5.59 在极小极大问题 (5.83) 上的具体化和改进, 其中包括 (5.83) 特有的新的互补松弛条件. 从而得到该定理的一个具体化和改进. 简单起见, 只在上导数条件 (5.79) 的情形下描述定理 5.59 在极小极大问题上的变体.

定理 5.62(极小极大问题的最优性条件) 设 $\bar{x}$ 是约束极小极大问题 (5.83) 的局部最优解, 其中 $\bar{z} := f(\bar{x})$, $f\colon X \to Z$ 是 Asplund 空间之间的映射, 在 $\bar{x}$ 附近相对于 Ω 是连续的. 假设 Ω 在 $\bar{x}$ 附近是局部闭的. 且存在 $z_0 \in Z$ 使得 $\langle z^*, z_0 \rangle = 1$ 对任意的 $z^* \in \Lambda$ 成立. 如果下述条件之一成立:

(a) 集合

$$\Theta := \left\{ z \in Z \,\middle|\, \langle z^*, z \rangle \leqslant 0, \ \forall z^* \in \Lambda \right\}$$

在 $f(\bar{x}) - \varphi(\bar{x}) z_0$ 处是 SNC 的.

(b) 逆映射 f_Ω^{-1} 在 $(\bar{z}, \bar{x})$ 处是 PSNC 的.

则存在 $\bar{z}^* \in Z^*$ 满足包含关系

$$0 \in D_N^* f_\Omega(\bar{x})(\bar{z}^*), \quad \text{其中} \quad \bar{z}^* \neq 0, \tag{5.84}$$

$$\bar{z}^* \in \text{cl}^* \text{co}[\text{cone}\Lambda] = \text{cl}^* \left\{ \sum_{i=1}^m \alpha_i z_i^* \,\middle|\, \alpha_i \geqslant 0, z_i^* \in \Lambda, m \in \mathbb{N} \right\} \tag{5.85}$$

和互补松弛条件

$$\left\langle \bar{z}^*, \bar{z} - \varphi(\bar{x}) z_0 \right\rangle = 0. \tag{5.86}$$

证明　在定理所给数据的基础上构造集合

$$\overline{\Theta} := \Theta + \big(\varphi(\bar{x})z_0 - \bar{z}\big),$$

并且证明 $\bar{x}$ 相对于 Ω 是局部 $(f, \overline{\Theta})$- 最优的. 由于 $\bar{z} - \varphi(\bar{x})z_0 \in \Theta$, 故有 $0 \in \overline{\Theta}$. 需验证当某 $z_k \to 0$ 时条件 (5.76) 成立. 假设相反, 取 $z_k := z_0/k, \forall k \in \mathbb{N}$, 从 $\bar{x}$ 的邻域 U 中找到 x, 使得 $x \in \Omega$ 且

$$\begin{aligned}\big\langle z^*, f(x)\big\rangle - \varphi(\bar{x}) &= \big\langle z^*, f(x)\big\rangle - \varphi(\bar{x})\big\langle z^*, z_0\big\rangle \\ &\leqslant -\big\langle z^*, z_0\big\rangle/k < 0, \quad \forall z^* \in \Lambda\end{aligned}$$

当 $k \to \infty$ 时成立, 这与极小极大问题 (5.84) 中的 $\bar{x}$ 的局部最优性矛盾. 在这种情形下应用定理 5.59(ii), 并且考虑到 $\overline{\Theta}$ 的凸性, 找到 $\bar{z}^*$ 满足 (5.84) 和

$$\big\langle \bar{z}^*, z - \big(\bar{z} - \varphi(\bar{x})z_0\big)\big\rangle \leqslant 0, \quad \forall\, z \in \Theta. \tag{5.87}$$

余下需证 (5.87) 蕴涵着 (它实际上等价于) (5.85) 和 (5.86) 这两个条件. 事实上, 由 (5.87) 和 Θ 的锥结构, 有不等式

$$\big\langle \bar{z}^*, \alpha z - (\bar{z} - \varphi(\bar{x})z_0)\big\rangle \leqslant 0, \quad \forall \alpha > 0, \quad z \in \Theta.$$

当 $\alpha \to \infty$ 时通过取极限有 $\langle \bar{z}^*, z\rangle \leqslant 0$ 对任意的 $z \in \Theta$ 成立, 从而 (5.85) 成立. 而且, 由 $\bar{z} - \varphi(\bar{x})z_0 \in \Theta$, 有不等式

$$\big\langle \bar{z}^*, \bar{z} - \varphi(\bar{x})z_0\big\rangle \leqslant 0$$

成立. 相反的不等式由 (5.87) 当 $z = 0$ 时即得. 因此得 (5.86), 这就完成了定理的证明. △

值得注意的是, 如果集合 Λ 由有限多线性无关的元素组成, 那么定理 5.62 中关于这个集合的所有假设自动成立. 此时一般极小极大问题 (5.83) 实际上简化为最小化有限个实值函数的最大值:

$$\min \varphi(x) = \max\big\{\varphi_i(x) \,\big|\, i = 1, \cdots, n\big\} \quad \text{s.t.} \quad x \in \Omega. \tag{5.88}$$

简单起见, 假设所有 φ_i 都是局部 Lipschitz 的, 下面给出关于问题 (5.88) 定理 5.62 的一个简单推论.

推论 5.63(关于有限个函数的极小极大问题)　设 $\bar{x}$ 是 Asplund 空间 X 中问题 (5.88) 的局部最优解, 这里所有的 $\varphi_i\colon X \to \mathbb{R}$ 在 $\bar{x}$ 附近都是 Lipschitz 连续的, Ω 在该点附近是局部闭的. 则存在实数 $\lambda_i \geqslant 0$, $i = 1, \cdots, n$, 使得

$$\lambda_i\big(\varphi_i(\bar{x}) - \varphi(\bar{x})\big) = 0, \quad i = 1, \cdots, n,$$

$$0 \in \partial \left(\sum_{i=1}^{n} \lambda_i \varphi_i + \delta(\cdot; \Omega) \right) (\bar{x}) \subset \partial \left(\sum_{i=1}^{n} \lambda_i \varphi_i \right) (\bar{x}) + N(\bar{x}; \Omega).$$

证明 由定理 5.62, 在 $Z = \mathbb{R}^m, \Lambda$ 由基本单位向量组成时直接可得. 注意到, 定理中具有最后一个包含关系的这些必要最优性条件也可利用关于最大值函数的次微分分析法则的定理 3.46(ii) 由标量最优化问题的必要最优性条件而得到. △

类似上面多目标结果, 根据定理 5.62 和推论 5.63, 可得具有算子和其他约束的极小极大问题的必要最优性条件, 以及仅具有不等式约束时的上次微分条件.

5.3.3 集值映射的极点原理

本节的下一个目标是得到约束多目标问题的必要最优性条件, 其中最优性的概念通过定义 5.55 意义下的闭序关系来描述. 正如在 5.3.1 小节中看到的那样, 这个概念可以不同于 5.3.2 小节中研究的广义序最优性. 从变分几何的观点来说, 一般闭序导致与非线性变形 (而非平移) 有关的集值映射 (而非集合) 的极点系统. 这一小节研究这样的集值映射的极点系统, 并且得到它们关于近似和确切两种形式的极点原理的合适版本. 出于对变形参数的依赖性考虑, 其中的确切形式需要在移动集合的情形下极限法向量和 SNC 性质的某些推广. 下面从集值映射系统的局部极性的定义开始.

定义 5.64(集值映射的极点系统) 设 $S_i\colon M_i \rightrightarrows X,\ i = 1, \cdots, n$ 是从度量空间 (M_i, d_i) 到 Banach 空间 X 的集值映射. 称 $\bar{x}$ 是系统 $\{S_1, \cdots, S_n\}$ 在 $(\bar{s}_1, \cdots, \bar{s}_n)$ 的局部极点, 如果 $\bar{x} \in S_1(\bar{s}_1) \cap \cdots \cap S_n(\bar{s}_n)$ 且存在 $\bar{x}$ 的邻域 U 使得 $\forall \varepsilon > 0$, 存在 $s_i \in \operatorname{dom} S_i$ 满足条件

$$d_i(s_i, \bar{s}_i) \leqslant \varepsilon, \quad \operatorname{dist}\big(\bar{x}; S_i(s_i)\big) \leqslant \varepsilon, \quad i = 1, \cdots, n,$$

$$S_1(s_1) \cap \cdots \cap S_n(s_n) \cap U = \varnothing. \tag{5.89}$$

此时 $\{S_1, \cdots, S_n, \bar{x}\}$ 称为在 $(\bar{s}_1, \cdots, \bar{s}_n)$ 的极点系统.

易见上面的定义将定义 2.1 中集合极性的概念推广到集值映射的情形. 事实上, 简单起见, 考虑两个集合的极点系统 $\{\Omega_1, \Omega_2, \bar{x}\}$, 通过令

$$M_1 := X, \quad M_2 := \{0\}, \quad S_1(s_1) := \Omega_1 + s_1, \quad S_2(0) := \Omega,$$

即可把这个集合极点系统化为上面集值映射的概念, 这对应于定义 5.64 中集值映射的线性性 (或集合的平移). 下面的例子表明关于集合变形的极点系统不能简化为通过它们的平移而得的极点系统. 事实上, 考虑如下定义

$$S_1(s_1) := \big\{(x, y) \in \mathbb{R}^2 \,\big|\, |x| - 2|y| \geqslant s_1\big\}, \tag{5.90}$$

$$S_2(s_2) := \{(x, y) \in \mathbb{R}^2 \mid |y| - 2|x| \geqslant s_2\}$$

的移动集合 (即集值映射), 它们可以看成是初始集合 $\Omega_1 := S_1(0)$ 与 $\Omega_2 := S_2(0)$ 的变形. 可验证 $(0,0)$ 是 $\{S_1, S_2\}$ 在定义 5.64 意义下的局部极点, 但在定义 2.1 的意义下却不是.

涉及集合变形的极点系统的主要例子与定义 5.55 描述的闭序关系的多目标最优化问题相关联.

例 5.65(具有闭序的多目标最优化的极点) 设 $f\colon X \to Z$ 是 Banach 空间之间的映射, $\prec$ 是 Z 上的闭序关系, 具有水平集 $\mathcal{L}(z)$, 设 $\bar{x}$ 是约束多目标问题

$$\min\ f(x) \quad \text{s.t.} \quad x \in \Omega$$

的最优解, 这里"min"理解为相对于序 $\prec$. 则 $(\bar{x}, f(\bar{x}))$ 是由

$$S_1(s_1) := \Omega \times \operatorname{cl}\mathcal{L}(s_1), \quad \text{其中} \quad M_1 := \mathcal{L}(f(\bar{x})) \cup \{f(\bar{x})\},$$

$$S_2(s_2) = S_2 := \{(x, f(x)) \mid x \in X\}, \quad \text{其中} \quad M_2 := \{0\}$$

定义的集值映射 $S_i\colon M_i \rightrightarrows X \times Z \quad (i = 1, 2)$ 的系统在 $(f(\bar{x}), 0)$ 的局部极点.

证明 首先注意到由 $\prec$ 的局部饱足性质得 $(\bar{x}, f(\bar{x})) \in S_1(f(\bar{x})) \cap S_2$. 为建立 (5.89), 假设相反, 根据序关系的非自反性, 对给定 $(\bar{x}, f(\bar{x}))$ 的任意邻域 U, 找到接近于 $f(\bar{x})$(但不等于 $f(\bar{x})$) 的点 $s_1 \in \mathcal{L}(f(\bar{x}))$, 使得

$$S_1(s_1) \cap S_2 \cap U \neq \varnothing.$$

这蕴涵着存在 $\bar{x}$ 附近的 x 满足 $(x, f(x)) \in S_1(s_1) = \Omega \times \operatorname{cl}\mathcal{L}(s_1)$. 因此根据 $\prec$ 的几乎可传递性, 有 $x \in \Omega$, $f(x) \prec f(\bar{x})$. 这与所考虑的约束多目标问题中 $\bar{x}$ 的局部最优性矛盾. △

在建立集值映射的极点原理及其在多目标最优化中的应用之前, 下面给出另外两个极点系统的例子, 它们当然也具有独立的意义.

例 5.66(二人对策中的极点) 设 $(\bar{x}, \bar{y}) \in \Omega \times \Theta$ 是在 Banach 空间中的子集 $\Omega \subset X$ 和 $\Theta \subset Y$ 上的支付函数 $\varphi\colon X \times Y \to \mathbb{R}$ 的鞍点, 即

$$\varphi(x, \bar{y}) \leqslant \varphi(\bar{x}, \bar{y}) \leqslant \varphi(\bar{x}, y), \quad \forall (x, y) \in \Omega \times \Theta.$$

定义集值映射 $S_1\colon \Big[\varphi(\bar{x}, \bar{y}), \infty\Big) \times \Big(-\infty, \varphi(\bar{x}, \bar{y})\Big] \rightrightarrows \Omega \times \mathbb{R} \times \Theta \times \mathbb{R}$ 和集合 $S_2 \subset \Omega \times \mathbb{R} \times \Theta \times \mathbb{R}$ 为

$$S_1(\alpha, \beta) := \Omega \times [\alpha, \infty) \times \Theta \times (-\infty, \beta], \quad S_2 := \operatorname{hypo}\varphi(\cdot, \bar{y}) \times \operatorname{epi}\varphi(\bar{x}; \cdot).$$

则 $(\bar{x},\varphi(\bar{x},\bar{y}),\bar{y},\varphi(\bar{x},\bar{y}))$ 是系统 $\{S_1,S_2\}$ 在 $(\varphi(\bar{x},\bar{y}),\varphi(\bar{x},\bar{y}))$ 的局部极点.

证明 显然有

$$\big(\bar{x},\varphi(\bar{x},\bar{y}),\bar{y},\varphi(\bar{x},\bar{y})\big)\in S_1\big(\varphi(\bar{x},\bar{y}),\varphi(\bar{x},\bar{y})\big)\cap S_2.$$

进一步, 由鞍点 $(\bar{x},\bar{y})$ 的定义得

$$S_1(\alpha,\beta)\cap S_2=\varnothing,\quad \forall(\alpha,\beta)\in\big[\varphi(\bar{x},\bar{y}),\infty\big)\times\big(-\infty,\varphi(\bar{x},\bar{y})\big],$$

而且 $(\alpha,\beta)\neq(\varphi(\bar{x},\bar{y}),\varphi(\bar{x},\bar{y}))$. 因此 $(\bar{x},\varphi(\bar{x},\bar{y}),\bar{y},\varphi(\bar{x},\bar{y}))$ 是 $\{S_1,S_2\}$ 如上意义下的局部极点. △

例 5.67(时间最优控制中的极点) 设 $\bar{\tau}$ 是下面最优控制问题的最优解:

最小化瞬态时间 $\tau>0$, 受约束于满足相应于可测控制 $u(\cdot)$ 的常微分方程

$$\dot{x}(t)=f\big(x(t),u(t)\big),\quad x(0)=x_0,\quad u(t)\in U,\qquad \text{a.e.}\quad t\in[0,\tau] \tag{5.91}$$

的绝对连续轨道 $x\colon[0,\tau]\to\mathbb{R}^n$ 上的端点约束 $x(\tau)=0$. 考虑由

$$S_1(s_1):=\big\{x(s_1)\in\mathbb{R}^n\,\big|\,x(\cdot)\ \text{在(5.91)中在}[0,s_1]\text{上是可行的}\big\}$$

定义的可达集集值映射 $S_1\colon(0,\infty)\rightrightarrows\mathbb{R}^n$, 并设 $S_2=\{0\}\subset\mathbb{R}^n$. 则 $0\in\mathbb{R}^n$ 是当 $M_1=(0,\infty), M_2=\{0\}\subset\mathbb{R}$ 时定义 5.64 意义下系统 $\{S_1,S_2\}$ 在 $(\bar{\tau},0)$ 的局部极点.

证明 由定义直接可得. △

接下来建立集值映射系统类似定理 2.20 的近似形式的极点原理. 这个结果实际上等价于定理 2.20 中集合系统的近似极点原理, 而且是 Asplund 空间的另一刻画.

定理 5.68(集值映射的近似极点原理) 设 $S_i\colon M_i\rightrightarrows X$ 是度量空间 (M_i,d_i) 到 Banach 空间 X 的集值映射, $i=1,\cdots,n$. 则下面的结论等价:

(a) X 是 Asplund 空间.

(b) 对任何在 $(\bar{s}_1,\cdots,\bar{s}_n)$ 的极点系统 $\{S_1,\cdots,S_n,\bar{x}\}$, 如果每个 S_i 在 $\bar{s}_i$ 附近是闭值的, 则近似极点原理成立. 即对任意的 $\varepsilon>0$, 存在 $s_i\in\operatorname{dom}S_i$, $x_i\in S_i(s_i)$ 和 $x_i^*\in X^*$, $i=1,\cdots,n$, 满足

$$d_i(s_i,\bar{s}_i)\leqslant\varepsilon,\quad \|x_i-\bar{x}\|\leqslant\varepsilon,\quad x_i^*\in\widehat{N}\big(x_i;S_i(s_i)\big)+\varepsilon\mathbb{B}^*, \tag{5.92}$$

$$x_1^*+\cdots+x_n^*=0,\quad \|x_1^*\|+\cdots+\|x_n^*\|=1. \tag{5.93}$$

(c) 对任何在 $(\bar{s}_1,\cdots,\bar{s}_n)$ 的极点系统 $\{S_1,\cdots,S_n,\bar{x}\}$, 如果每个 S_i 在 $\bar{s}_i$ 附近是闭值的, 则近似极点原理的 ε- 法向量形式成立, 即在 (b) 中的近似极点原理中, 在 (5.92) 中用 $\widehat{N}_\varepsilon\big(x_i;S_i(s_i)\big)$ 代替 $\widehat{N}\big(x_i;S_i(s_i)\big)+\varepsilon\mathbb{B}^*$, $i=1,\cdots,n$.

证明 首先注意到由于总有

$$\widehat{N}(\bar{x};\Omega)+\varepsilon\mathbb{B}^*\subset\widehat{N}_\varepsilon(\bar{x};\Omega),$$

所以 (b)⇒(c). 进一步注意到 (c) 中集值映射的 ε- 极点原理蕴涵着定义 2.5(i) 中集合系统的相应结果. 从而定理中的蕴涵关系 (c)⇒(a) 由定理 2.20 中的 (c)⇒(a) 可得. 余下需证 (a)⇒(b), 即近似极点原理对任意 Asplund 空间中集值映射的极点系统成立. 类似 2.2 节的步骤, 基于 Fréchet 光滑空间中直接的变分讨论及可分约化方法即可得证. 下面, 利用定理 2.26(i) 中的 Ekeland 变分原理和引理 2.32 中半 - Lipschitz 和的最小值点的模糊次梯度条件 (它等价于集合系统的近似极点原理), 给出另一证明.

设 $\bar{x}$ 是系统 $\{S_1,\cdots,S_n\}$ 在 $(\bar{s}_1,\cdots,\bar{s}_n)$ 的局部极点, 这里 X 是 Asplund 空间. 取 $U:=\bar{x}+r\mathbb{B}$, 对任给的 $\varepsilon>0$, 选取 $\varepsilon'>0$ 满足

$$\varepsilon'<\min\left\{\varepsilon^2/(5\varepsilon+12n^2+\varepsilon^2),r^2/4\right\}.$$

则相应于 ε' 在定义 5.64 中取 $s_1,\cdots,s_n$. 记 $\Omega:=S_1(s_1)\times\cdots\times S_n(s_n)$, 并构造函数

$$\varphi(y_1,\cdots,y_n):=\sum_{i,j=1}^n\|y_i-y_j\|+\delta\Big((y_1,\cdots,y_n);\Omega\Big),$$

其中 $(y_1,\cdots,y_n)\in U^n$. 该函数是 l.s.c. 的, 并且在完备度量空间 U^n 上是正的. 对满足

$$\|y_i'-y_j'\|\leqslant\operatorname{dist}\big(\bar{x};S_i(s_i)\big)+\operatorname{dist}\big(\bar{x};S_j(s_j)\big)+\varepsilon'\leqslant3\varepsilon'$$

的任意 $y_i'\in S_i(s_i)$, 有 $\varphi(y_1',\cdots,y_n')\leqslant3n^2\varepsilon'<\varepsilon^2/4$. 把定理 2.26(i) 中的 Ekeland 变分原理应用于上面的函数 φ, 找到 $x_i'\in y_i'+(\varepsilon/2)\mathbb{B}\subset\bar{x}+\varepsilon\mathbb{B}$, $i=1,\cdots,n$, 使得扰动函数

$$\sum_{i,j=1}^n\Big\|y_i-y_j\Big\|+\frac{\varepsilon}{2}\sum_{i=1}^n\Big\|y_i-x_i'\Big\|+\delta\big((y_1,\cdots,y_n);\Omega\big)\tag{5.94}$$

在 $(x_1',\cdots,x_n')$ 处在 U^n 上达到它的全局最小值. 不失一般性, 假设 $U^n=X^n$, 并且记

$$\psi(y_1,\cdots,y_n):=\sum_{i,j=1}^n\Big\|y_i-y_j\Big\|,\quad(y_1,\cdots,y_n)\in X^n.$$

根据构造可知 $\psi(x_1',\cdots,x_n')>0$. 现对 (5.94) 应用定理 2.20 和引理 2.32(i), 并且考虑到

$$\widehat{\partial}\delta\Big((y_1,\cdots,y_n);\Omega\Big)=\widehat{N}\big(y_1;S_1(s_1)\big)\times\cdots\times\widehat{N}\big(y_n;S_n(s_n)\big),\quad\forall y_i\in S_i(s_i),$$

找到 $x_i \in S_i(s_i)\cap(x_i'+\varepsilon'\mathbb{B}) \subset (\bar{x}+\varepsilon\mathbb{B})$, $z_i \in x_i'+\varepsilon'\mathbb{B}$, $i=1,\cdots,n$ 和 $(-x_1^*,\cdots,-x_n^*) \in \widehat{\partial}\psi(z_1,\cdots,z_n)$, 使得

$$0 \in (-x_1^*,\cdots,-x_n^*) + \widehat{N}\big(x_1;S_1(s_1)\big) \times \cdots \times \widehat{N}\big(x_n;S_n(s_n)\big) + \varepsilon'(n+1)(\mathbb{B}^*)^n.$$

上面的关系显然蕴涵着对选取的数 ε, 有

$$x_i^* \in \widehat{N}\big(x_i;S_i(s_i)\big) + \varepsilon\mathbb{B}^*$$

对任意的 $i=1,\cdots,n$ 成立, 从而 (5.92) 成立.

现证 $x_1^*,\cdots,x_n^*$ 也满足 (5.93). 如果必要进一步缩小 ε', 可使 $\psi(z_1,\cdots,z_n)>0$. 注意到由包含关系 $(-x_1^*,\cdots,-x_n^*) \in \widehat{\partial}\psi(z_1,\cdots,z_n)$ 可推出

$$\begin{aligned}
&\langle -x_1^*-\cdots-x_n^*, h\rangle \\
&\leqslant \liminf_{t\to 0}\frac{\psi(z_1+th,\cdots,z_n+th)-\psi(z_1,\cdots,z_n)}{t} \\
&= \liminf_{t\to 0}\frac{\displaystyle\sum_{i,j=1}^n \|(z_i+th)-(z_j+th)\| - \sum_{i,j=1}^n \|z_i-z_j\|}{t} = 0
\end{aligned}$$

对任意单位向量 $h \in X$ 成立. 这就得到 (5.93) 中的第一关系式 (Euler 方程). 余下需证

$$\|x_1^*\|+\cdots+\|x_n^*\| \geqslant 1,$$

由规范化它蕴涵着 (5.93) 中的第二个关系式. 为此, 注意到函数 ψ 是正齐次的, 因此包含关系 $(-x_1^*,\cdots,-x_n^*) \in \widehat{\partial}\psi(z_1,\cdots,z_n)$ 蕴涵着

$$\begin{aligned}
\sum_{i=1}^n \langle -x_i^*, -z_i\rangle &\leqslant \liminf_{t\to 0}\frac{\psi(z_1-tz_1,\cdots,z_n-tz_n)-\psi(z_1,\cdots,z_n)}{t} \\
&= -\psi(z_1,\cdots,z_n).
\end{aligned}$$

在 Euler 方程中应用 $-x_i^* = x_2^*+\cdots+x_n^*$, 有

$$\begin{aligned}
\psi(z_1,\cdots,z_n) &\leqslant \sum_{i=1}^n \langle -x_i^*, z_i\rangle = \sum_{i=2}^n \langle x_i^*, z_1-z_i\rangle \\
&\leqslant \max\{\|x_i^*\| \mid i=2,\cdots,n\}\sum_{i=2}^n \|z_1-z_i\| \\
&\leqslant \max\{\|x_i^*\| \mid i=1,\cdots,n\}\psi(z_1,\cdots,z_n).
\end{aligned}$$

由于 $\psi(z_1,\cdots,z_n)>0$, 上式给出了估计

$$\max\left\{\|x_i^*\|\,\middle|\,i=1,\cdots,n\right\}\geqslant 1,$$

这就完成了定理的证明. △

下一个目标是得到集值映射确切/极限形式的极点原理, 类似定理 2.22 中对集合系统的确切极点原理. 通过对近似极点原理中的关系式 (5.92) 和 (5.93) 在 $\varepsilon\downarrow 0$ 时取极限得到这样的结果是很自然的. 然而, 这里稍微不同于集合极点原理的情形, 由于现在 (5.92) 中的集合 $S_i(s_i)$ 是移动的, 即它们依赖于当 $\varepsilon\downarrow 0$ 时收敛于 $\bar{s}_i$ 的某些点. 为执行极限过程并得到极点原理的适当极限形式, 需要描述移动集合的极限法向量, 也需添加适当的法紧性条件, 从而可以在无限维情形取极限. 首先定义移动集合的极限法向量所构成的锥, 这个锥在有限和无限维空间中都是有用的.

定义 5.69(移动集合的极限法向量) 设 $S\colon Z\rightrightarrows X$ 是从度量空间 Z 到 Banach 空间 X 的集值映射, $(\bar{z},\bar{x})\in\operatorname{gph}S$. 则

$$N_+\big(\bar{x},S(\bar{z})\big):=\mathop{\mathrm{Limsup}}_{(z,x)\xrightarrow{\operatorname{gph}S}(\bar{z},\bar{x}),\varepsilon\downarrow 0}\widehat{N}_\varepsilon\big(x;S(z)\big)\tag{5.95}$$

称为 $S(\bar{z})$ 在 $\bar{x}$ 的推广法锥. 如果

$$N_+\big(\bar{x},S(\bar{z})\big)=N\big(\bar{x},S(\bar{z})\big),\tag{5.96}$$

则称映射 S 在 $(\bar{z},\bar{x})$ 是法向半连续的.

值得注意的是, 如果 X 是 Asplund 空间, S 在 $\bar{x}$ 附近是闭值的, 则可在 (5.95) 中等价地令 $\varepsilon=0$. 这由表达式 (2.51) 所给的 Asplund 空间中 ε- 法向量的表示直接可得. 还注意到 (5.95) 中的正规性概念与集值映射 $S(\cdot)$ 的 (广义) 可微性无关: 变量 z 只是移动集合的参数, 在极限过程中才会用到.

(5.96) 中的包含关系"$\supset$"总成立, 即, 关于移动集合 $S(\cdot)$ 在 (5.95) 的极限过程中可能出现的极限法向量显然比只考虑集合 $S(\bar{x})$ 在 (1.2) 的极限过程中出现的极限法向量要多. 然而, 当 $z\to\bar{z}$ 时集合 $S(z)$ 表现比较好时, $N_+(\bar{x},S(\bar{z}))$ 与基本法锥 $N(\bar{x},S(\bar{z}))$ 是可以相同的, 不仅仅在它们是参数独立时. 下面给出性质 (5.96) 的一个简单的充分条件; 这个方向的更多结果也可参阅本章的评注.

命题 5.70(移动集合的法向半连续性) 设 $S\colon Z\rightrightarrows X$ 是从度量空间 Z 到 Banach 空间 X 的集值映射. 则在下面两种情形下, S 在 $(\bar{z},\bar{x})\in\operatorname{gph}S$ 处是法向半连续的:

(i) 在 $\bar{z}$ 附近 $S(z)=g(z)+\Omega$, 这里 $\Omega\subset X$ 是任一非空集合, $g\colon Z\to X$ 在 $\bar{z}$ 是连续的.

(ii) S 在 $\bar{z}$ 附近是凸值的, 且在该点是内半连续的, 即

$$S(\bar{z})\subset\liminf_{z\to\bar{z}}S(z).$$

证明 在情形 (i) 下法向半连续性由定义 (1.2)、(5.95) 及 $g(\cdot)$ 连续性直接可得. 值得注意的是, 这种情形对涉及固定集合平移的确切极点原理的应用来说是足够的.

现在考虑情形 (ii). 取 $x^* \in N_+(\bar{x}, S(\bar{z}))$, 找到序列 $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$, $z_k \to \bar{z}$ 和 $x_k^* \xrightarrow{w^*} x^*$, 使得

$$x_k \in S(z_k), \quad x_k^* \in \widehat{N}_{\varepsilon_k}\big(x_k; S(z_k)\big), \quad \forall k \in \mathbb{N}.$$

利用关于凸集的 ε- 法向量表示的命题 1.3, 得显式表述

$$\langle x_k^*, u - x_k \rangle \leqslant \varepsilon_k \|u - x_k\|, \quad \forall u \in S(z_k).$$

下面证明 (ii) 中的内半连续性假设蕴涵着

$$\langle x^*, u - \bar{x} \rangle \leqslant 0, \quad \forall u \in S(\bar{z}),$$

由于对凸集而言基本法锥与凸分析中的法锥相同, 从而有 $x^* \in N(\bar{x}; S(\bar{z}))$. 事实上, 假设相反, 上面的蕴涵关系在某点 $\bar{u} \in S(\bar{z})$ 不成立, 即 $\langle x^*, \bar{u} - \bar{x} \rangle > 0$. 利用 S 在 $\bar{z}$ 的内半连续性, 对给定的 $\bar{u}$ 和序列 $z_k \to \bar{z}$, 找到序列 $u_k \to \bar{u}$, 使得 $u_k \in S(z_k)$ 对任意的 $k \in \mathbb{N}$ 成立. 则有表示

$$\langle x_k^*, u_k - x_k \rangle = \langle x^*, \bar{u} - \bar{x} \rangle + \Big[\langle x_k^* - x^*, \bar{u} - \bar{x} \rangle + \langle x_k^*, u_k - \bar{u} \rangle - \langle x_k^*, x_k - \bar{x} \rangle\Big].$$

易见由 x_k, u_k, x_k^* 的相应收敛性及 $\{x_k^*\}$ 的有界性知, 方括号中的所有项当 $k \to \infty$ 时都趋于 0. 这就得

$$\langle x_k^*, u_k - x_k \rangle > \varepsilon_k \|u_k - x_k\|$$

对大的 $k \in \mathbb{N}$ 成立, 这与 ε- 法向量的上述表达式矛盾, 从而完成了命题的证明. △

为进一步得到无限维像空间情形下集值映射的确切极点原理, 需要下面的集值映射的法紧性, 这个性质涉及集值映射的像而不是它们的图, 这与基本的 SNC 是不同的.

定义 5.71(移动集合的 SNC 性质) 称度量空间 Z 和 Banach 空间 X 之间的集值映射 $S\colon Z \rightrightarrows X$ 在 $(\bar{z}, \bar{x}) \in \operatorname{gph} S$ 是像 SNC(或 ISNC) 的, 如果对任意满足

$$x_k^* \in \widehat{N}_{\varepsilon_k}\big(x_k; S(z_k)\big), \quad \varepsilon_k \downarrow 0, \quad (z_k, x_k) \xrightarrow{\operatorname{gph} S} (\bar{z}, \bar{x}), \quad x_k^* \xrightarrow{w^*} 0$$

的序列 $(\varepsilon_k, z_k, x_k, x_k^*)$ 有 $\|x_k^*\| \to 0 \ (k \to \infty)$.

若 X 是有限维的, 则该性质当然自动成立. 若 $g\colon Z \to X$ 在 $\bar{z}$ 是连续的, $\Omega \subset X$ 在 $\bar{x} - g(\bar{z})$ 是 SNC 的, 则当 S 在 $\bar{z}$ 附近有表示

$$S(z) = g(z) + \Omega$$

时, 这个性质也成立. 如果 X 是 Asplund 空间, S 在 $\bar{z}$ 附近是闭值的, 则可在定义 5.71 中等价地令 $\varepsilon_k = 0$. 与固定集合的情形类似, 对移动集合来说, 上面的 ISNC 性质与 CEL 性质的相应结果之间有深刻的关系. 特别地, 如果存在数 $\alpha, \eta > 0$ 和紧集 $C \subset X$, 使得

$$\widehat{N}_\varepsilon\big(x; S(z)\big) \subset \left\{x^* \in X^* \,\middle|\, \eta\|x^*\| \leqslant \varepsilon\alpha + \max_{c \in C} \big|\langle x^*, c\rangle\big|\right\}$$

对任意 $(z, x) \in \operatorname{gph} S \cap \Big((\bar{z}, \bar{x}) + \eta\mathbb{B}_{Z\times X}\Big)$ 成立, 则 Banach 空间之间的映射 $S: Z \rightrightarrows X$ 在 $(\bar{z}, \bar{x})$ 是 ISNC 的. 如果 S 在 $(\bar{z}, \bar{x})$ 是一致 CEL 的, 即存在紧集 $C \subset Z$, $(\bar{z}, \bar{x})$ 的邻域 $V \times U$ 和 Z 中原点的邻域 O 及数 $\gamma > 0$, 使得

$$S(x) \cap U + tO \subset S(x) + \gamma C, \quad \forall x \in U, \quad t \in (0, \gamma),$$

则前面的关于 $\widehat{N}_\varepsilon\big(x; S(z)\big)$ 的包含关系当然成立, 请比照定理 1.26 证明. 与定义 1.24 一致, S 被称为在 $(\bar{x}, \bar{z})$ 附近是一致上图 Lipschitz 的, 如果 C 可选为单点集. 如果存在 $\bar{z}$ 的邻域 V, 使得对 $z \in V, S(z)$ 是凸的且 $\operatorname{int}\big(\cap_{z\in V} S(z)\big) \neq \varnothing$, 则对任意 $\bar{x} \in S(\bar{z})$ 这个条件总成立; 请比照命题 1.25 的证明. 与 1.2.5 小节类似, 可定义集值映射的部分 ISNC 性质, 并得出对一致类 Lipschitz 和部分 CEL 集值映射来说这个性质成立.

值得一提的是, 推广法锥 (5.95) 和定义 5.71 中的 ISNC 性质以及它们的映射/函数对应物和部分类似物也拥有完备的分析法则, 与本书所建立的基本结构和 SNC 性质的相应分析法则类似. 下面没有给出这样的结构及其应用; 它们的表述和证明与“非移动”对象的情形类似.

现在已准备好来建立集值映射系统的确切/极限极点原理, 这个极点原理推广了 (实际上等价于) 在定理 2.22 中所得到的关于集合系统的确切极点原理.

定理 5.72(集值映射的确切极点原理)

(i) 设 $S_i\colon M_i \rightrightarrows X$, $i = 1, \cdots, n$, 是从度量空间 (M_i, d_i) 到 Asplund 空间 X 的集值映射. 假设 $\bar{x}$ 是系统 $\{S_1, \cdots, S_n\}$ 在 $(\bar{s}_1, \cdots, \bar{s}_n)$ 的局部极点, 其中每个 S_i 在 $\bar{s}_i$ 附近是闭的, 而且除了一个可能的例外, 所有的 S_i 在它们图的相应点 $(\bar{s}_i, \bar{x})$ 处是 ISNC 的. 则下面的确切极点原理成立:

存在

$$x_i^* \in N_+\big(\bar{x}; S_i(\bar{s}_i)\big), \quad i = 1, \cdots, n$$

满足广义 Euler 方程

$$x_1^* + \cdots + x_n^* = 0, \quad 且 \quad (x_1^*, \cdots, x_n^*) \neq 0.$$

(ii) 反之, 设确切极点原理对任意具有像空间 X 的两个集值映射的极点系统 $\{S_1, S_2, \bar{x}\}$ 都成立, 这里, 两个映射 S_i 在相应点 $\bar{s}_i$ 附近都是闭值的, 而且其中之一在 $(\bar{s}_i, \bar{x})$ 是 ISNC 的. 则 X 是 Asplund 空间.

证明 (ii) 由定理 2.22(ii) 直接可得, 这是由于关于集值映射系统的确切极点原理蕴涵着关于集合系统的确切极点原理, 而当集合是固定集合时, 关于移动集合的 ISNC 性质简化为标准的 SNC 性质. 余下证 (i) 成立.

为此在 X 是 Asplund 空间时应用定理 5.68(b) 中给出的近似极点原理. 它保证对任意的 $k \in \mathbb{N}$, 存在 s_{ik} 使得 $d(s_{ik}, \bar{s}_i) \leqslant \dfrac{1}{k}$, $x_{ik} \in \bar{x} + \dfrac{1}{k}\mathbb{B}$, 及 $x_{ik}^* \in \widehat{N}\big(x_{ik}; S_i(s_{ik})\big)$, $i = 1, \cdots, n$, 满足关系

$$1 + \frac{1}{k} \geqslant \|x_{1k}^*\| + \cdots + \|x_{nk}^*\| \geqslant 1 - 1/k, \quad \|x_{1k}^* + \cdots + x_{nk}^*\| \leqslant 1/k. \tag{5.97}$$

由于根据 X 的 Asplund 性质对偶球 $\mathbb{B}^* \subset X^*$ 是弱 * 列紧的, 且 $\{x_{ik}^*\}$ 是有界的, 可找到 $x_i^* \in X^*$, 使得沿 $k \to \infty$ 的子序列有 $x_{i_k}^* \xrightarrow{w^*} x_i^*$ $(i = 1, \cdots, n)$. 现取当 $k \to \infty$ 时的极限, 并且应用定义 (5.95), 得定理中想要的等式关系, 余下需证 $(x_1^*, \cdots, x_n^*)$ 的非平凡性.

假设所有的 $x_i^* = 0$. 为确定起见, 假设前 $n - 1$ 个映射 S_i 在 $(\bar{s}_i, \bar{x})$ 是 ISNC 的 $(i = 1, \cdots, n - 1)$. 则由 x_i^* 的选取有 $\|x_{ik}^*\| \to 0$ $(k \to \infty)$, $i = 1, \cdots, n - 1$. 对 (5.97) 中的第二个关系式取极限, 也得 $\|x_{nk}^*\| \to 0$. 对大的 $k \in \mathbb{N}$, 这显然与 (5.97) 中的第一个关系式矛盾, 从而完成了定理的证明. △

值得注意的是, 在定理 5.72 中推广法锥 (5.95) 一般来说不能由基本法锥 (1.2) 代替, 除非相应的映射 S_i 为法向半连续的. 事实上, 考虑在 (5.90) 中定义的集值映射极点系统 $\{S_1, S_2\}$, 其在 $(\bar{s}_1, \bar{s}_2) = (0, 0)$ 处的局部极点为 $\bar{x} = 0 \in \mathbb{R}^2$. 易验证 S_1 和 S_2 在 $(0, 0, 0)$ 点都不是法向半连续的, 而且

$$N\big(0; S_1(0)\big) \cap \big[-N\big(0; S_2(0)\big)\big] = \{0\}.$$

因此当 N_+ 由 N 代替时, 定理 5.72 中的类似结果对集值映射的这个极点系统不成立.

5.3.4 相对于闭序的最优性条件

本小节给出推广极点原理在一般约束多目标最优化问题上的一些应用, 这里目标函数相对于闭序关系是"最小化的". 首先考虑下面只有几何约束的多目标问题:

$$\min f(x) \quad (\text{相对于} \prec) \quad \text{s.t.} \quad x \in \Omega, \tag{5.98}$$

这里, $f\colon X \to Z$ 是 Banach 空间之间的映射, $\Omega \subset X$, $\prec$ 是 Z 上具有定义 5.55 中

的移动水平集 $\mathcal{L}(\cdot)$ 的非自反序关系. 下面的定理给出了 (5.98) 关于近似/模糊和确切/极限形式的必要最优性条件.

定理 5.73(具有闭序关系和几何约束的问题的最优性条件)　设 $\bar{x}$ 是问题 (5.98) 的局部最优解, 其中 $\bar{z} := f(\bar{x})$, 序关系 $\prec$ 是闭的, 空间 X 和 Z 都是 Asplund 空间. 假设 f 在 $\bar{x}$ 附近是连续的, Ω 在该点附近是局部闭的. 则下面的结论成立:

(i) 对任意的 $\varepsilon > 0$, 存在 $(x_0, x_1, z_0, z_1, x^*, z^*) \in X^2 \times Z^2 \times X^* \times Z^*$, 满足 $x_0, x_1 \in \bar{x} + \varepsilon\mathbb{B}_X$, $z_0, z_1 \in \bar{z} + \varepsilon\mathbb{B}_Z$,

$$x^* \in \widehat{N}(x; \Omega), \quad z^* \in \widehat{N}\big(z_1; \mathrm{cl}\,\mathcal{L}(z_0)\big),$$

满足 $\|(x^*, z^*)\| = 1$, 而且

$$0 \in x^* + \widehat{D}^* f(x_0)(z^*) + \varepsilon\mathbb{B}_{X^*}.$$

进一步, 如果 f 在 $\bar{x}$ 附近是 Lipschitz 连续的, 则有

$$0 \in \widehat{\partial}\langle z^*, f\rangle(x_0) + \widehat{N}(x_1; \Omega) + \varepsilon\mathbb{B}_{X^*}, \quad 且 \quad \|z^*\| = 1.$$

(ii) 假设 f 在 $\bar{x}$ 是 SNC 的, 或 Ω 在 $\bar{x}$ 是 SNC 的且序 $\prec$ 的水平集产生的映射 $\mathrm{cl}\mathcal{L}\colon Z \rightrightarrows Z$ 在 $(\bar{z}, \bar{z})$ 是 ISNC 的. 则存在不全为 0 的 x^* 和 z^* 满足

$$x^* \in D_N^* f(\bar{x})(z^*) \cap \Big(-N(\bar{x}; \Omega)\Big), \quad z^* \in N_+\Big(\bar{z}; \mathrm{cl}\mathcal{L}(\bar{z})\Big).$$

进一步, 如果 f 在 $\bar{x}$ 是严格 Lipschitz 的, 并且或者 $\dim Z < \infty$, 或者 Ω 在 $\bar{x}$ 是 SNC 的及 $\mathrm{cl}\mathcal{L}$ 在 $(\bar{z}, \bar{z})$ 是 ISNC 的, 则有

$$0 \in \partial\langle z^*, f\rangle(\bar{x}) + N(\bar{x}; \Omega), \quad z^* \in N_+\big(\bar{z}; \mathrm{cl}\mathcal{L}(z)\big) \setminus \{0\}.$$

证明　先根据定理 5.68 中的近似极点原理证 (i) 成立. 在例 5.65 中已证明 $(\bar{x}, \bar{z})$ 是系统 $\{S_1, S_2\}$ 在 $(\bar{z}, 0)$ 的局部极点, 其中 $S_i\colon M_i \rightrightarrows X \times Z \quad (i = 1, 2)$ 定义为

$$S_1(z) = \Omega \times \mathrm{cl}\mathcal{L}(z), \quad S_2 \equiv \mathrm{gph}\, f.$$

由于空间 $X \times Z$ 是 Asplund 空间, S_i 在所给的一般假设下都是局部闭值的, 应用定理 5.68(b) 中的结论, 得 $z_0 \in \bar{z} + \varepsilon\mathbb{B}_Z$, $(x_i, z_i) \in (\bar{x}, \bar{z}) + \varepsilon\mathbb{B}_{X\times Z}, (i = 1, 2)$ 满足

$$(x_1^*, z_1^*) \in \widehat{N}\big((x_1, z_1); S_1(z_0)\big), \quad (x_2^*, z_2^*) \in \widehat{N}\big((x_2, z_2); S_2\big),$$

$$\big\|(x_1^*, z_1^*) + (x_2^*, z_2^*)\big\| \leqslant \varepsilon, \quad \|(x_1^*, z_1^*)\| + \|(x_2^*, z_2^*)\| \geqslant 1 - \varepsilon,$$

其中 $x_1 \in \Omega$, $z_1 \in \mathrm{cl}\mathcal{L}(z_0)$, $z_2 = f(x_2)$. 考虑到 S_1, S_2 的结构和命题 1.2 中关于 $\widehat{N}$ 的乘积公式, 由上面关系中的第一行得

$$x_1^* \in \widehat{N}(x_1; \Omega), \quad z_1^* \in \widehat{N}\big(z_1; \mathrm{cl}\mathcal{L}(z_0)\big), \quad x_2^* \in \widehat{D}^* f(x_2)(-z_2^*).$$

置 $x_0 := x_2$, $x^* := x_1^*$, $z^* := z_1^*$, 并利用规范化来保证 $\|(x^*, z^*)\| = 1$. 于是利用上面关系中的第二行, 如果必要缩小 ε, 当 f 在 $\bar{x}$ 附近仅假定为连续时, 易得 (x^*, z^*) 满足 (i) 中的所有结论. 如果 f 在该点附近是 Lipschitz 连续的, 则

$$\widehat{D}^* f(x_0)(z^*) = \widehat{\partial}\langle z^*, f\rangle(x_0),$$

从而完成了结论 (i) 的证明.

为证 (ii), 对所考虑的极点系统 $\{S_1, S_2, (\bar{x}, \bar{z})\}$ 应用定理 5.72(i) 中的确切极点原理. S_i 的结构和命题 1.2 中的乘积公式确保定理的 ISNC/SNC 假设蕴涵着定理 5.72 中所需的 ISNC 性质, 还有

$$N_+\big((\bar{x}, \bar{z}); S_1(\bar{z})\big) = N(\bar{x}; \Omega) \times N_+\big(\bar{z}; \mathrm{cl}\mathcal{L}(\bar{z})\big),$$

$$N_+\big((\bar{x}, \bar{z}); S_2\big) = \big\{(x^*, z^*) \mid x^* \in D_N^* f(\bar{x})(-z^*)\big\}.$$

这样 (ii) 的第一部分中的所有结论由定理 5.72 中的确切极点原理直接可得.

为证 (ii) 中第二部分的必要最优性条件成立, 根据定理 3.28, 只需注意到当 f 在 $\bar{x}$ 是严格 Lipschitz 的, 有

$$D_N^* f(\bar{x})(z^*) = \partial\langle z^*, f\rangle(\bar{x}),$$

以及如果 f 在 $\bar{x}$ 附近是 Lipschitz 连续的, 而 $\dim Z < \infty$, 则它在 $\bar{x}$ 是 SNC 的; 参见推论 1.69(i). 这就完成了定理的证明. $\triangle$

值得一提的是, 当 $f\colon X \to Z$ 在 $\bar{x}$ 是严格 Lipschitz 的, 而且 X 是 Asplund 空间, 则 f 在 $\bar{x}$ 的 SNC 性质根据推论 3.30 等价于 Z 的有限维性. 还注意到定理 5.73(ii) 中映射 $\mathrm{cl}\mathcal{L}$ 在 $(\bar{z}, \bar{z})$ 的 ISNC 性质和水平集映射 $\mathcal{L}$ 的相应性质的唯一差别在于 $\bar{z} \in \mathrm{cl}\mathcal{L}(\bar{z})$, 然而 $\bar{z} \notin \mathcal{L}(\bar{z})$, 由于序关系 $\prec$ 是局部饱足和非自反的.

注 5.74(多目标问题的最优性条件的比较) 以上建立了关于具有几何约束的多目标最优化问题的必要最优性条件的两个基本结果: 定理 5.59 和定理 5.73. 尽管在这两个定理中所考虑的多目标最优性的两个概念都推广了大多数传统概念, 但一般来说它们是不同的; 参见 5.3.1 小节中的结果及讨论. 然而, 在定理 5.59 和定理 5.73 所得的必要最优性条件中有许多相同之处. 特别地, 比较其结论 (ii) 中的上导数条件. 对 $f_\Omega(x) = f(x) + \Delta(x; \Omega)$ 利用命题 3.12 中的上导数和法则, 并利用规范条件

$$D_N^* f(\bar{x})(0) \cap \big(-N(\bar{x}; \Omega)\big) = \{0\},$$

如果 f 在 $\bar{x}$ 是 SNC 的, 或者 Ω 在 $\bar{x}$ 是 SNC 的且 Θ 在 0 是 SNC 的, 则由 (5.79) 及加在定理 5.59(ii) 中的法紧性条件, 可得相对于 Ω 的 $\bar{x}$ 的 (f,Θ)- 最优性蕴涵着存在 $(x^*,z^*)\neq 0$ 满足

$$0\in x^*+D_N^*f(\bar{x})(z^*),\quad x^*\in N(\bar{x};\Omega),\quad z^*\in N(0;\Theta).$$

在一般的情形下 (即使在有限维空间中), 定理 5.59(ii) 给出了关于广义序最优性的更精细的必要条件. 另一方面, 定理 5.73(ii) 适用于相对于闭序关系的多目标最优化问题, 这不能由极点系统中固定集合的传统平移来处理, 而是涉及移动集合的非线性形变.

类似 5.3.2 小节中广义序最优性的情形及本章前面的结果, 可得到具有闭序关系的多目标问题在算子和泛函约束下的定理 5.73 的各种推论. 所有这些推论都基于第 3 章中建立的广泛的广义微分和 SNC 分析法则的应用. 作为这些结果的一个例子, 下面给出定理 5.73(ii) 中的上导数最优条件在算子约束多目标问题上的推论.

推论 5.75(具有闭序关系和算子约束的问题的最优性条件) 设 $\prec$ 是 Z 上的闭序关系且有水平集 $\mathcal{L}(\cdot)$, 设 $\bar{x}$ 是多目标最优化问题:

$$\min f(x)\quad (\text{相对于 }\prec)\quad \text{s.t.}\quad x\in G^{-1}(\Lambda)$$

的局部最优解, 这里, $f\colon X\to Z$, $G\colon X\rightrightarrows Y$ 是 Asplund 空间之间的映射, 其中 $\bar{z}:=f(\bar{x})$, $\Lambda\subset Y$. 假设 f 是连续的, $S(\cdot):=G(\cdot)\cap\Lambda$ 在 $\bar{x}$ 附近是内半紧的, 集合 $\operatorname{gph}G$ 和 Λ 在相应点附近是局部闭的. 则存在 x^* 和 z^*, 不全为 0, 使得

$$-x^*\in D_N^*f(\bar{x})(z^*),\quad z^*\in N_+(\bar{z};\operatorname{cl}\mathcal{L}(\bar{z})),$$

$$x^*\in\bigcup\Big[D_N^*G(\bar{x},\bar{y})(y^*)\Big|y^*\in N(\bar{y};\Lambda),\bar{y}\in S(\bar{x})\Big]$$

成立, 其中假设关于 (f,G,Λ) 的下述条件之一成立:

(a) f 在 $\bar{x}$ 是 SNC 的, 规范条件

$$N(\bar{y};\Lambda)\cap\ker\widetilde{D}_M^*G(\bar{x},\bar{y})=\{0\}$$

对任意的 $\bar{y}\in S(\bar{x})$ 满足, 而且或者 G^{-1} 在 $(\bar{y},\bar{x})$ 是 PSNC 的, 或者对任意 $\bar{y}\in S(\bar{x})$, 映射 Λ 在 $\bar{y}$ 是 SNC 的.

(b) $\operatorname{cl}\mathcal{L}$ 在 $(\bar{z},\bar{z})$ 是 ISNC 的, 规范条件

$$N(\bar{y};\Lambda)\cap\ker D_N^*G(\bar{x},\bar{y})=\{0\}$$

对任意的 $\bar{y}\in S(\bar{x})$ 满足, 而且或者 G 在 $(\bar{x},\bar{y})$ 是 PSNC 的, 且 Λ 在 $\bar{y}$ 是 SNC 的, 或者对任意的 $\bar{y}\in S(\bar{x})$ 映射 G 在 $(\bar{x},\bar{y})$ 是 SNC 的.

证明 为从定理 5.73(ii) 的上导数最优性条件得此推论, 只需在 (a) 的假设下应用定理 3.8 给出 $G^{-1}(\Lambda)$ 的基本法锥的表示, 及在 (b) 的假设下应用定理 3.84 确保 $G^{-1}(\Lambda)$ 的 SNC 性质. △

接下来考虑等式和不等式型的泛函约束下相对于闭序关系的多目标问题. 类似 5.3.2 小节, 可得两种类型的必要最优性条件, 分别涉及约束函数的基本下次梯度以及针对不等式约束的 Fréchet 上次梯度. 为简单起见, 只给出关于具有不等式约束问题的结果, 因为只有这些约束使得下和上次微分条件有区别.

定理 5.76(具有不等式约束的多目标问题的下和上次微分条件) 设 $\prec$ 是 Z 上的闭序关系, 且有水平集 $\mathcal{L}(\cdot)$, 设 $\bar{x}$ 是多目标问题:

$$\min f(x) \quad (\text{相对于 } \prec) \quad \text{s.t.} \quad \varphi_i(x) \leqslant 0, \quad i = 1, \cdots, m$$

的局部最优解, 这里 $f\colon X \to Z$ 在 $\bar{x}$ 附近是连续的, 其中 $\bar{z} := f(\bar{x})$, 同时 $\varphi_i\colon X \to \overline{\mathbb{R}}$ 仅在 $\bar{x}$ 有限 $(i = 1, \cdots, m)$. 假设 f 在 $\bar{x}$ 是 SNC 的, 或者 $\mathrm{cl}\mathcal{L}$ 在 $(\bar{z}, \bar{z})$ 是 ISNC 的. 下面的结论成立:

(i) 假设空间 X 和 Z 都是 Asplund 空间, 假设每个 φ_i 在 $\bar{x}$ 附近都是 Lipschitz 连续的. 则存在 $z^* \in Z^*$ 和乘子 $(\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m$ 满足

$$z^* \in N_+(\bar{x}; \mathrm{cl}\mathcal{L}(\bar{z})), \quad \lambda_i \geqslant 0, \quad \lambda_i \varphi_i(\bar{x}) = 0, \quad i = 1, \cdots, m \tag{5.99}$$

使得 $(z^*, \lambda_1, \cdots, \lambda_m) \neq 0$ 并且

$$0 \in D_N^* f(\bar{x})(z^*) + \partial\left(\sum_{i=1}^m \lambda_i \varphi_i\right)(\bar{x}) \subset D_N^* f(\bar{x})(z^*) + \sum_{i=1}^m \lambda_i \partial\varphi_i(\bar{x})$$

成立.

(ii) 假设 Z 是 Asplund 空间, 同时 X 有一个 Lipschitz $\mathcal{C}^1$ 阻尼函数 (当 X 有一个 Fréchet 光滑重赋范时这自动成立). 则对任意 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$, $i = 1, \cdots, m$, 存在 $0 \neq (z^*, \lambda_1, \cdots, \lambda_m) \in Z^* \times \mathbb{R}^m$, 满足 (5.99) 和

$$0 \in D_N^* f(\bar{x})(z^*) + \sum_{i=1}^m \lambda_i x_i^*.$$

证明 定理的结论 (i) 中下次微分最优性条件由推论 5.73(i) 和 (ii) 直接可得, 其中 $Y = \mathbb{R}^m, G(x) = (\varphi_1(x), \cdots, \varphi_m(x)), \Lambda = \mathbb{R}^m_-$. 事实上, 只需注意到此时有

$$N(\bar{y}, \Lambda) = \left\{(\lambda_1, \cdots, \lambda_m) \in \mathbb{R}^m \,\middle|\, \lambda_i \geqslant 0, \lambda_i \varphi_i(\bar{x}) = 0, i = 1, \cdots, m\right\},$$

$$D^*G(\bar{x})(\lambda_1, \cdots, \lambda_m) = \partial\left(\sum_{i=1}^m \lambda_i \varphi_i\right)(\bar{x}) \subset \sum_{i=1}^m \lambda_i \partial\varphi_i(\bar{x}).$$

为证 (ii) 中的上次微分条件成立, 任取 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$ $(i = 1, \cdots, m)$, 根据定理 1.88(ii) 中 Fréchet 次梯度的变分描述, 找到在 $\bar{x}$ 的某邻域 U 内连续可微的函数 $s_i\colon X \to \overline{\mathbb{R}}$, 使得

$$s_i(\bar{x}) = \varphi_i(\bar{x}), \quad \nabla s_i(\bar{x}) = x_i^*, \quad s_i(x) \geqslant \varphi_i(x), \quad i = 1, \cdots, m.$$

易见 $\bar{x}$ 是多目标问题

$$\min f(x) \quad (\text{相对于 } \prec) \quad \text{s.t.} \quad s_i(x) \leqslant 0, i = 1, \cdots, m$$

的局部最优解. 现对这个问题应用结论 (i) 中的最优性条件, 就完成了 (ii) 的证明. △

这一小节的最后简要讨论 (推广的) 极点原理在一类具有多个局中人/参与者的多目标博弈中的一些应用. 这类问题可以粗略地描述为具有几个局中人的博弈, 其中每个局中人都想从空间 X_i 中选择一个策略 $\bar{x}_i$, 使得它们 $\prec_i$ 最小化 (相对于 Y 上的序关系 $\prec_i$) 在所有其他局中人的选择 $\bar{x}_j, i \neq j$ 给定时的目标映射 $f: X_1 \times \cdots \times X_n \to Z$.

这是一个一般的博弈情形, 特别地, 包括每一个局中人都可有不同的目标映射 $f_i\colon X_1 \times \cdots \times X_n \to Z_i$ 的情形. 在这种情形下, 有 $f := (f_1, \cdots, f_n)\colon X_1 \times \cdots \times X_n \to Z := Z_1 \times \cdots \times Z_n$, 其中, Z 上的序 $\prec_i$ 定义为: $z \prec_i v$, $z, v \in Z$, 如果 $z_i \prec_i v_i$, $z_i, v_i \in Z_i$.

众所周知, 在所有博弈论中的一个基本概念就是鞍点概念. 下面给出这个概念关于上面多目标情形的一个广义版本, 这里 $\prec$ 代表 $(\prec_1, \cdots, \prec_n)$.

定义 5.77(多目标博弈的鞍点) 称点 $\bar{x} = (\bar{x}_1, \cdots, \bar{x}_n)$ 是 $f\colon X_1 \times \cdots \times X_n \to Z$ 的局部 $\prec$- 鞍点, 如果对每个 $i = 1, \cdots, n$, 都存在 $\bar{x}_i$ 的邻域 U_i, 使得

$$f(\bar{x}) \prec_i f(\bar{x}_1, \cdots, \bar{x}_{i-1}, x_i, \bar{x}_{i+1}, \cdots, \bar{x}_n)$$

对任意的 $x_i \in U_i$ 成立.

注意到这个鞍点概念可以不同于例 5.66 中考虑的具有不依赖于局中人和空间序关系的通常概念. 事实上, 设支付映射 $f\colon \mathbb{R}^4 \to \mathbb{R}^2$ 由

$$f(x, y, u, v) := (x^2 + u, -y^2 - e^v)$$

给出. 将变量分组使得 x 和 y 属于第一个局中人, u 和 v 属于第二个局中人. 这意味着 $X_1 = X_2 = Z = \mathbb{R}^2$. 对第一个局中人 $Z = \mathbb{R}^2$ 上的序 $\prec_1$ 定义为 $(w, s) \prec_1 (\widetilde{w}, \widetilde{s})$, 如果 $w < \widetilde{w}$, $s \geqslant \widetilde{s}$ 或者 $w \leqslant \widetilde{w}, s > \widetilde{s}$. 对第二个局中人 $Z = \mathbb{R}^2$ 上的序 $\prec_2$ 定义为 $(w, s) \prec_2 (\widetilde{w}, \widetilde{s})$, 如果 $w < \widetilde{w}$ 和 $s < \widetilde{s}$. 这是 Pareto 和弱 Pareto 最优性的混合. 可验证形式为 $(0, 0, u, v)$ 的任何点总是这些序下的局部 $\prec$- 鞍点.

现在给出具有局中人策略的其他约束的多目标博弈的必要最优性条件. 简单起见只描述关于几何约束情形的结果.

给定如定义 5.77 中的 $f\colon X_1\times\cdots\times X_n\to Z$ 和 $\prec_i$, 以及约束集合 $\Omega_i\subset X_i$ $(i=1,\cdots,n)$, 考虑下面的多目标约束博弈 $\mathcal{G}$: 找到受约束于约束 $x_i\in\Omega_i\subset X_i$ $(i=1,\cdots,n)$ 的 f 的局部 $\prec$- 鞍点. 设 $\bar{x}$ 是博弈 $\mathcal{G}$ 的局部最优解. 根据 $\prec$- 鞍点的定义 5.77, 对任意局中人 i, $\bar{x}$ 的第 i 个分量 $\bar{x}_i$ 是下面多目标约束最优化问题:

$$\min f(\bar{x}_1,\cdots,\bar{x}_{i-1},x_i,\bar{x}_{i+1},\cdots,\bar{x}_n)\quad \text{s.t.}\quad x_i\in\Omega_i$$

是局部解. 这里, “最小化”理解为相对于 Z 上的序关系 $\prec_i$.

记 $f_i(x_i):=f(\bar{x}_1,\cdots,\bar{x}_{i-1},x_i,\bar{x}_{i+1},\cdots,\bar{x}_n)$, $\bar{z}_i:=f_i(\bar{x}_i)$, $i=1,\cdots,n$, 考虑由 Z 上的序关系 $\prec_i$ 诱导的水平集 $\mathcal{L}_i(z)$. 利用上面关于多目标最优化问题的结果, 根据集值映射极点原理的近似和确切版本, 下面得到多目标博弈中的必要最优性条件. 简洁起见, 只描述关于 Lipschitz 目标映射情形的相应结果.

定理 5.78(多目标博弈的最优性条件) 设 $\bar{x}=(\bar{x}_1,\cdots,\bar{x}_n)$ 是上面博弈 $\mathcal{G}$ 的局部最优解, 这里序关系 $\prec_i$ 在 Z 上是闭的, $X_1,\cdots,X_n,Z$ 都是 Asplund 空间. 假设映射 $f\colon X_1\times\cdots\times X_n\to Z$ 在 $\bar{x}$ 附近是 Lipschitz 连续的, 集合 $\Omega_i\subset X_i$ 在 $\bar{x}_i$ 附近是局部闭的 $(i=1,\cdots,n)$. 则下面的断言成立:

(i) 对任意的 $\varepsilon>0$, 存在 $(x_i,u_i,z_i,v_i,z_i^*)\in X_i\times X_i\times Z\times Z\times Z^*$ 满足 $x_i,u_i\in\bar{x}+\varepsilon\mathbb{B}_{X_i}, z_i,\ v_i\in\bar{z}_i+\varepsilon\mathbb{B}_Z$, $\|z_i^*\|=1$ 且

$$0\in\widehat{\partial}\langle z_i^*,f_i\rangle(x_i)+\widehat{N}(u_i;\Omega_i)+\varepsilon\mathbb{B}_{X_i^*},\quad z_i^*\in\widehat{N}\big(v_i;\mathrm{cl}\mathcal{L}_i(z_i)\big),\quad i=1,\cdots,n.$$

(ii) 假设 f 在 $\bar{x}$ 是严格 Lipschitz 的, 而且或者 $\dim Z<\infty$, 或者 Ω_i 在 $\bar{x}_i$ 是 SNC 的且 $\mathrm{cl}\mathcal{L}_i$ 在 $(\bar{z}_i,\bar{z}_i)$ 是 ISNC 的 $(i=1,\cdots,n)$. 则存在 $z_1^*,\cdots,z_n^*\in Z^*$ 使得 $\|z_i^*\|=1$,

$$0\in\partial\langle z_i^*,f_i\rangle(\bar{x}_i)+N(\bar{x}_i;\Omega_i),\quad z_i^*\in N_+\big(\bar{z}_i;\mathrm{cl}\mathcal{L}_i(\bar{z}_i)\big),\quad i=1,\cdots,n$$

成立.

证明 由于对每个局中人 $i=1,\cdots,n$, $\bar{x}$ 的第 i 个分量 $\bar{x}_i$ 是上面所描述的多目标最优化问题的局部最优解, 把定理 5.73 中的两个断言应用于这些问题, 就得 (i) 和 (ii) 中的必要最优性条件. △

5.3.5 具有均衡约束的多目标最优化

本小节主要研究包括由参数变分系统控制的具有形式为

$$0\in q(x,y)+Q(x,y)$$

的均衡约束的多目标最优化问题. 5.3 节已经在具有单 (实值) 目标函数的 MPEC 的框架下研究过这样的约束. 现研究具有均衡约束的一类多目标最优化问题, 其中的最优解在定义 5.53 给出的广义序最优性或者在定义 5.55 给出的闭序关系的意义下来理解. 正如在 5.3.1 小节中讨论的那样, 这些多目标概念都包含与 Pareto 型最优性/有效性及类似概念相关的标准均衡概念. 因此下面研究的多目标最优化问题包括具有均衡约束的均衡问题 (EPEC), 它们对许多应用来说非常重要. 注意到, 关于多目标问题的上水平的均衡概念可由向量变分不等式描述; 特别地, 参见文献 [504] 及其中的参考文献. 方便起见, 本小节考虑的所有均衡型约束的多目标问题将采用简写符号 EPEC(或 EPEC 问题, 稍微有些滥用语言).

尽管 EPEC 除均衡约束外可能还有其他类型的约束 (几何、算子、泛函), 简洁起见, 这里都不考虑; 类似 5.2 节, 这些约束的结果也可以建立起来. 下面主要关注在参考最优解处描述的 EPEC 的点基/确切必要最优性条件.

首先研究 EPEC, 这里, 最优解应理解为在定义 5.53 中的广义序最优性意义下的最优解. 下面的结果给出了由一般依赖参数的集值映射描述的均衡约束的问题的抽象版本的必要最优性条件. 在它的描述中, 应用的 $F\colon X \rightrightarrows Y$ 在 $(\bar{x},\bar{y}) \in \operatorname{gph} F$ 的强 PSNC 性质与定义 1.67 和定义 3.3 一致, 即对任何满足

$$\varepsilon_k \downarrow 0, \quad (x_k, y_k) \to (\bar{x},\bar{y}), \quad x_k^* \in \widehat{D}^*_{\varepsilon_k} F(x_k,y_k)(y_k^*), \quad (x_k^*, y_k^*) \xrightarrow{w^*} (0,0)$$

的序列 $(\varepsilon_k, x_k, y_k, x_k^*, y_k^*) \in [0,\infty) \times (\operatorname{gph} F) \times X^* \times Y^*$, 有 $\|x_k^*\| \to 0\ (k \to \infty)$. 特别地, 这对在 $(\bar{x},\bar{y})$ 附近是部分 CEL 的映射 $F\colon X \rightrightarrows Y$ 成立; 参见定理 1.75. 值得注意的是, 在上面的关系式中对 Asplund 空间之间的闭图映射可等价地令 $\varepsilon_k = 0$.

定理 5.79(抽象 EPEC 的广义序最优性) 设 $f\colon X \times Y \to Z$, $\Theta \subset Z$ 且 $0 \in \Theta$, $S\colon X \rightrightarrows Y$ 且 $(\bar{x},\bar{y}) \in \operatorname{gph} S$. 假设 X, Y 是 Asplund 空间, 点 $(\bar{x},\bar{y})$ 是受约束于 $y \in S(x)$ 的局部 (f,Θ)- 最优解. 则下面的断言成立:

(i) 假设集合

$$\mathcal{E}(f,S,\Theta) := \Big\{(x,y,z) \in X \times Y \times Z \Big| f(x,y) - z \in \Theta, y \in S(x)\Big\}$$

在 $(\bar{x},\bar{y},\bar{z})$ 附近是局部闭的, 其中 $\bar{z} := f(\bar{x},\bar{y})$, 而且 $\dim Z < \infty$. 则存在 $z^* \in Z^*$ 满足

$$(0, -z^*) \in N\Big((\bar{x},\bar{y},\bar{z}); \mathcal{E}(f,S,\Theta)\Big), \quad z^* \in N(0;\Theta) \setminus \{0\}.$$

(ii) 假设 Z 是 Asplund 空间, f 在 $(\bar{x},\bar{y})$ 附近是连续的, $\operatorname{gph} S$ 和 Θ 分别在 $(\bar{x},\bar{y})$ 和 0 附近是局部闭的. 考虑下面的两种情形:

(a) Θ 在 0 是 SNC 的,

$$\Big[(x^*, y^*) \in D^*_M f(\bar{x},\bar{y})(0),\ -x^* \in D^*_N S(\bar{x},\bar{y})(y^*)\Big] \Rightarrow x^* = y^* = 0,$$

并且或者 S 在 $(\bar{x},\bar{y})$ 是 SNC 的, 或者 f 在该点是 PSNC 的; 当 f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 则 f 在 $(\bar{x},\bar{y})$ 的 PSNC 性质及上面的规范条件自动成立.

(b) f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, f^{-1} 在 $(\bar{z},\bar{x},\bar{y})$ 是强 PSNC 的, 并且

$$\left[(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(0),\ -x^*\in D_N^*S(\bar{x},\bar{y})(y^*)\right]\Rightarrow x^*=y^*=0.$$

则在上述任一情形中都存在 $z^*\in N(0;\Theta)\setminus\{0\}$ 满足

$$0\in D_N^*f(\bar{x},\bar{y})(z^*)+N\big((\bar{x},\bar{y});\operatorname{gph}S\big),\tag{5.100}$$

而且, 如果 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 则 (5.100) 等价于

$$0\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+N\big((\bar{x},\bar{y});\operatorname{gph}S\big).$$

证明 注意到 EPEC 问题等价于在几何约束 $(x,y)\in\Omega:=\operatorname{gph}S$ 下定理 5.59 研究的关于两个变量的映射 f 的多目标最优化问题. 因此断言 (i) 由定理 5.59(i) 直接可得. 为证 (ii), 当 Θ 在 0 是 SNC 的或 $(f+\Delta(\cdot;\operatorname{gph}S))^{-1}$ 在 $(\bar{z},\bar{x},\bar{y})$ 是 PSNC 的时, 应用定理 5.59(ii), 则存在 $z^*\in N(0;\Theta)\setminus\{0\}$ 满足

$$0\in D_N^*\big(f+\Delta(\cdot;\operatorname{gph}S)\big)(\bar{x},\bar{y}).$$

接下来, 对特殊的和 $f+\Delta(\cdot;\operatorname{gph}S)$ 应用命题 3.12 中的上导数和法则, 在那个命题的极限规范条件 (3.25) 下, 有

$$0\in D_N^*f(\bar{x},\bar{y})(z^*)+N((\bar{x},\bar{y});\operatorname{gph}S).$$

此式自动成立, 如果

$$D_M^*f(\bar{x},\bar{y})(0)\cap\Big(-N\big((\bar{x},\bar{y});\operatorname{gph}S\big)\Big)=\{0\},$$

且 f 在 $(\bar{x},\bar{y})$ 是 PSNC 的, 或 S 在该点是 SNC 的; 如果 f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 这当然成立. 因此在 (a) 的假设下得 (5.100).

为证在情形 (b) 下 (5.100) 成立, 只需验证由 (b) 中的假设可得 $\big(f+\Delta(\cdot;\operatorname{gph}S)\big)^{-1}$ 在 $(\bar{z},\bar{x},\bar{y})$ 是 PSNC 的. 即对满足

$$(x_k^*,y_k^*)\in\widehat{D}^*\big(f+\Delta(\cdot;\operatorname{gph}S)\big)(x_k,y_k)(z_k^*),\quad \|(x_k^*,y_k^*)\|\to 0,\quad z_k^*\xrightarrow{w^*}0$$

的任意序列 $(x_k,y_k,x_k^*,y_k^*,z_k^*)$ 且 $(x_k,y_k)\to(\bar{x},\bar{y})$, 应有 $\|z_k^*\|\to 0\ (k\to\infty)$. 由定理 3.10 的证明可得, (b) 中的规范条件蕴涵着上面考虑的关于 Fréchet 上导数

$\widehat{D}^*\big(f+\Delta(\cdot;\mathrm{gph}\,S)\big)(x_k,y_k)$ 的模糊和法则成立, 从而保证存在 $\varepsilon_k\downarrow 0$, $(x_{ik},y_{ik})\to(\bar{x},\bar{y})$ $(i=1,2)$ 和 $(\widetilde{x}_k^*,\widetilde{y}_k^*,\widetilde{z}_k^*)$, 使得

$$(\widetilde{x}_k^*,\widetilde{y}_k^*)\in\widehat{D}^*f(x_{1k},y_{1k})(\widetilde{z}_k^*)+\widehat{N}\Big((x_{2k},y_{2k});\mathrm{gph}S\Big)$$

和 $\|(\widetilde{x}_k^*,\widetilde{y}_k^*,\widetilde{z}_k^*)-(x_k^*,y_k^*,z_k^*)\|\leqslant\varepsilon_k$ 对任意的 $k\in\mathbb{N}$ 成立. 因此对某

$$(x_{1k}^*,y_{1k}^*)\in\widehat{D}^*f(x_{1k},y_{1k})(\widetilde{z}_k^*)\quad 和\quad (x_{2k}^*,y_{2k}^*)\in\widehat{N}\Big((x_{2k},y_{2k});\mathrm{gph}S\Big),$$

有 $(\widetilde{x}_k^*,\widetilde{y}_k^*)=(x_{1k}^*,y_{1k}^*)+(x_{2k}^*,y_{2k}^*)$. 由于 f 在 $(\bar{x},\bar{y})$ 附近是局部 Lipschitz 的, 所以序列 (x_{1k}^*,y_{1k}^*) 在 $X^*\times Y^*$ 中有界; 因此, 根据 $X\times Y$ 的 Asplund 性质, 它包含弱 * 收敛于某 $(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(0)$ 的子序列. 根据 $\|(\widetilde{x}_k^*,\widetilde{y}_k^*)\|\to 0,(x_{2k}^*,y_{2k}^*)=(\widetilde{x}_k^*,\widetilde{y}_k^*)-(x_{1k}^*,y_{1k}^*)$, 沿 $k\to 0$ 的子序列有 $(x_{2k}^*,y_{2k}^*)\xrightarrow{w^*}(-x^*,-y^*)\in N\big((\bar{x},\bar{y});\mathrm{gph}S\big)$. 根据 (b) 中的规范条件得 $x^*=y^*=0$. 这蕴涵着 $(x_{1k}^*,y_{1k}^*,z_k^*)\xrightarrow{w^*}0$, 其中 $(x_{1k}^*,y_{1k}^*)\in\widehat{D}^*f(x_{1k},y_{1k})(z_k^*)$. 利用 f^{-1} 在 $(\bar{z},\bar{x},\bar{y})$ 的强 PSNC 性质, 最后得 $\|z_k^*\|\to 0$. 定理的最后一个结论由定理 3.28 的标量化公式可得. △

定理 5.79 中所得的关于抽象 EPEC 的必要最优性条件在一般约束规范条件下是以标准/规范形式给出的. 下面的推论给出了非规范 (Fritz John) 形式的必要最优性条件, 此时对原始数据不需加规范条件.

推论 5.80(抽象 EPEC 的非规范条件)　设 $(\bar{x},\bar{y})$ 在约束 $y\in S(x)$ 下是局部 (f,Θ)- 最优的, 这里, $f\colon X\times Y\to Z$, $\Theta\subset Z$, $S\colon X\rightrightarrows Y$ 满足定理 5.79(ii) 中的假设. 若下列情形之一成立:

(a) f 在 $(\bar{x},\bar{y})$ 是 PSNC 的, Θ 在 0 是 SNC 的;

(b) S 和 Θ 分别在 $(\bar{x},\bar{y})$ 和 0 是 SNC 的;

(c) f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, f^{-1} 在 $(\bar{z},\bar{x},\bar{y})$ 是强 PSNC 的, 其中 $\bar{z}=f(\bar{x},\bar{y})$. 则存在 $0\neq(x^*,y^*,z^*)\in X^*\times Y^*\times Z^*$, 使得必要最优性条件

$$(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*),\quad -x^*\in D_N^*S(\bar{x},\bar{y})(y^*),\quad z^*\in N(0;\Theta)$$

成立.

证明　如果定理 5.79(ii) 的情形 (a) 或 (b) 中的规范条件成立, 那么有推论的最优性条件对 $z^*\neq 0$ 成立. 若这些约束规范不成立, 则直接蕴涵着想要的最优性条件在 $(x^*,y^*)\neq 0$ 时成立. △

下面建立由参数独立的广义方程/变分系统控制的均衡约束的多目标问题的必要最优性条件. 它们对应于

$$S(x):=\big\{y\in Y\,\big|\,0\in q(x,y)+Q(x,y)\big\}\tag{5.101}$$

时上面的抽象框架. 为得到 (5.101) 型具有均衡/变分约束的 EPEC 的最优性条件, 需对 (5.101) 中给出的映射 $S(\cdot)$ 应用定理 5.79 和推论 5.80 的结果. 简单起见, 下面只给出这样问题的那些不需要约束规范的最优性条件, 即与推论 5.80 相似. 标准/规范形式的最优性条件通过定理 5.79 类似 5.2.2 小节中关于 MPEC 的下次微分条件可得.

定理 5.81(由变分系统控制的 EPEC 的广义序最优性) 设 $f\colon X\times Y\to Z$ 是 Asplund 空间之间的映射且 $\bar z:=f(\bar x,\bar y)$, 设 $\Theta\subset Z$ 且 $0\in\Theta$, 设 $(\bar x,\bar y)$ 是在约束

$$0\in q(x,y)+Q(x,y)$$

下局部 (f,Θ)- 最优的, 这里 $q\colon X\times Y\to P$, $Q\colon X\times Y\rightrightarrows P$ 是映入 Asplund 空间 P 的映射, 且记 $\bar p:=-q(\bar x,\bar y)$. 假设 f 和 q 在 $(\bar x,\bar y)$ 附近是连续的, Θ 在 0 附近是闭的, Q 在 $(\bar x,\bar y,\bar p)$ 附近是闭图的. 考虑下面的情形:

(a) f 在 $(\bar x,\bar y)$ 是 PSNC 的, Θ 和 Q 分别在 0 和 $(\bar x,\bar y,\bar p)$ 是 SNC 的, 且有规范条件

$$\Big[(x^*,y^*)\in D_N^*q(\bar x,\bar y)(p^*)\cap\big(-D_N^*Q(\bar x,\bar y,\bar p)(p^*)\big)\Big]\Rightarrow x^*=y^*=p^*=0,$$

当 q 在 $(\bar x,\bar y)$ 是严格 Lipschitz 的, 这等价于

$$\Big[0\in\partial\langle p^*,q\rangle(\bar x,\bar y)+D_N^*(\bar x,\bar y,\bar p)(p^*)\Big]\Rightarrow p^*=0. \tag{5.102}$$

(b) f 在 $(\bar x,\bar y)$ 是 PSNC 的, Θ 在 0 是 SNC 的, $\dim P<\infty$, q 在 $(\bar x,\bar y)$ 附近是 Lipschitz 连续的, 而且 (5.102) 成立.

(c) Θ 在 0 是 SNC 的, q 在 $(\bar x,\bar y)$ 是 PSNC 的, 而且 (a) 中的规范条件满足.

(d) f 在 $(\bar x,\bar y)$ 附近是 Lipschitz 连续的, f^{-1} 在 $(\bar z,\bar x,\bar y)$ 是强 PSNC 的, Q 在 $(\bar x,\bar y,\bar p)$ 是 SNC 的, 而且 (a) 中的规范条件成立.

(e) f 和 g 在 $(\bar x,\bar y)$ 附近是 Lipschitz 连续的, $\dim P<\infty$, f^{-1} 在 $(\bar z,\bar x,\bar y)$ 是强 PSNC 的, 而且 (5.102) 满足.

在上面任一情形下, 都存在 $(x^*,y^*,z^*,p^*)\in X^*\times Y^*\times Z^*\times P^*$ 满足 $(x^*,y^*,z^*)\neq 0$, $z^*\in N(0;\Theta)$, 且

$$(x^*,y^*)\in D_N^*f(\bar x,\bar y)(z^*)\cap\Big(-D_N^*q(\bar x,\bar y)(p^*)-D_N^*Q(\bar x,\bar y,\bar p)(p^*)\Big).$$

进一步, 如果 f 在 $(\bar x,\bar y)$ 是严格 Lipschitz 连续的, 则上面的最优性条件可等价地写为如下形式: 存在 $z^*\in N(0;\Theta)\setminus\{0\}$ 和 $p^*\in P^*$ 满足

$$0\in\partial\langle z^*,f\rangle(\bar x,\bar y)+D_N^*q(\bar x,\bar y)(p^*)+D_N^*Q(\bar x,\bar y,\bar p)(p^*).$$

证明 根据推论 5.80 中的最优性条件, 其中 S 由 (5.101) 定义, 需得出有效的条件使得 S 的上导数 $D_N^*S(\bar{x},\bar{y})$ 及其 SNC 性质可用 (5.101) 中的初始数据 (q,Q) 有效地表达. 为此, 首先应用定理 4.46, 它给出了在所给的条件下由 q 和 Q 的上导数表达的关于 S 的上上导数估计 (4.63). 这些假设与推论 5.80(a) 和 (c) 中的假设结合, 得到 (a), (b), (d) 及 (e) 情形下定理的结论. 余下考虑推论 5.80 中的情形 (c), 这需要关于 (5.101) 中的映射 S 的 SNC 性质的有效条件. 此时利用定理 4.59 的证明, 其中证明了 (基于定理 3.84) 映射 (5.101) 在 $(\bar{x},\bar{y})$ 是 SNC 的, 如果 q 在该点是 PSNC 的, 外加 (a) 中的规范条件和 Q 在 $(\bar{x},\bar{y},\bar{p})$ 的 SNC 性质. 这些假设与推论 5.80(b) 中的假设结合, 就证明了情形 (c) 下的结果成立. 定理的最后一个结论与往常一样由定理 3.28 中的标量化公式即得. △

与 4.4 小节及 5.2 小节中考虑的情形类似, 可得定理 5.81 的许多推论. 下面给出其中由复合次微分形式给出的均衡约束有关的一些结果, 它们对应用来说最有意思. 第一个结果涉及由具有复合势函数的所谓半变分不等式控制的均衡约束的 EPEC.

推论 5.82(由有复合势的 HVI 控制的 EPEC 的最优性条件) 设 $f\colon X\times Y\to Z$ 是连续映射且 $\bar{z}:=f(\bar{x},\bar{y})$, $\Theta\subset Z$ 是闭集且 $0\in\Theta$, $(\bar{x},\bar{y})$ 是均衡约束

$$0\in q(x,y)+\partial(\psi\circ g)(y)$$

下局部 (f,Θ)- 最优的, 这里 $q\colon X\times Y\to Y^*$, $g\colon Y\to W$, $\psi\colon W\to\overline{\mathbb{R}}$. 假定 W 是 Banach 空间, X 和 Z 是 Asplund 空间, $\dim Y<\infty$, 并且假设下面的条件成立:

(a) 或者 f 在 $(\bar{x},\bar{y})$ 是 PSNC 的, 且 Θ 在 0 是 SNC 的, 或者 f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 且 f^{-1} 在 $(\bar{z},\bar{x},\bar{y})$ 是强 PSNC 的.

(b) q 在 $(\bar{x},\bar{y})$ 是严格可微的且具有满射偏导数 $\nabla_x q(\bar{x},\bar{y})$.

(c) g 在 $\bar{y}$ 附近是连续可微的且具有满射导数 $\nabla g(\bar{y})$, 从 Y 到 $\mathcal{L}(Y,W)$ 的映射 $\nabla g(\cdot)$ 在 $\bar{y}$ 是严格可微的.

(d) $\partial\psi$ 的图在 $(\bar{w},\bar{v})$ 附近是局部闭的, 这里 $\bar{w}:=g(\bar{y})$, 而且 $\bar{v}\in W^*$ 是满足关系

$$-q(\bar{x},\bar{y})=\nabla g(\bar{y})^*\bar{v},\quad \bar{v}\in\psi(\bar{w})$$

的唯一泛函.

则存在 $(y^*,z^*,u)\in Y^*\times Z^*\times Y$ 满足 $(y^*,z^*)\neq 0$, $z^*\in N(0;\Theta)$, $(-\nabla_x q(\bar{x},\bar{y})^*u,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)$, 并且

$$-y^*\in\nabla_y q(\bar{x},\bar{y})^*u+\nabla^2\langle\bar{v},g\rangle(\bar{y})^*u+\nabla g(\bar{y})^*\partial_N^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{y})u\big).$$

进一步, 如果 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 则上面的最优性条件可写为如下形

式: 存在 $z^* \in N(0;\Theta)\setminus\{0\}$, $u \in Y$ 使得

$$0\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+\nabla q(\bar{x},\bar{y})^*u$$
$$+\Big(0,\nabla^2\langle\bar{v},g\rangle(\bar{y})^*u+\nabla g(\bar{y})^*\partial_N^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{y})u\big)\Big).$$

证明 在定理 5.81 中令 $Q(y)=\partial(\psi\circ g)(y)$, 利用定理 1.127 中的二阶次微分链式法则, 通过计算

$$D^*Q(\bar{y},\bar{p})(u)=\partial^2(\psi\circ g)(\bar{y},\bar{p})(u),\quad 其中\ \bar{p}:=-q(\bar{x},\bar{y})$$

可得. 注意到, 由于 $\nabla_x q(\bar{x},\bar{y})$ 是满射的且 Q 不依赖于参数 x, 由 $\dim Y<\infty$ 知 Q 是 SNC 的, 从而定理 5.81 的规范条件自动成立. △

下一个推论给出了 EPEC 的必要最优性条件, 其中的均衡约束由具有复合势的依赖参数的变分系统 (称为广义变分不等式 -GVI) 给出. 简单起见, 只考虑有限维中的顺从势 (函数) 的情形. 注意关于导数没有加满射性的假设.

推论 5.83(由具有顺从势控制的 EPEC 的广义序最优性) 设 $f: X\times Y\to Z$ 是连续映射, 设 $\Theta\subset Z$ 是闭集且 $0\in\Theta$, $(\bar{x},\bar{y})$ 在参数独立的均衡约束

$$0\in q(x,y)+\partial(\psi\circ g)(\bar{x},\bar{y})$$

下是局部 (f,Θ)- 最优的, 这里 $q\colon X\times Y\to X^*\times Y^*$, $g\colon X\times Y\to W$, $\psi\colon W\to\overline{\mathbb{R}}$, $\dim(X\times Y\times W)<\infty$, 而且 Z 是 Asplund 空间. 假设 q 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 势 $\varphi:=\psi\circ g$ 在该点是强顺从的. 记 $\bar{p}:=-q(\bar{x},\bar{y})\in\partial(\psi\circ g)$, $\bar{w}:=g(\bar{x},\bar{y})$,

$$M(\bar{x},\bar{y}):=\big\{\bar{v}\in W^*\,\big|\,\bar{v}\in\partial\psi(\bar{w}),\nabla g(\bar{x},\bar{y})^*\bar{v}=\bar{p}\big\},$$

并假设二阶规范条件

$$\partial^2\psi(\bar{w},\bar{v})(0)\cap\ker\nabla g(\bar{x},\bar{y})^*=\{0\},\quad\forall\bar{v}\in M(\bar{x},\bar{y}),$$
$$\Big[0\in\partial\langle u,q\rangle(\bar{x},\bar{y})+\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\big[\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})(u)$$
$$+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\big]\Big]\Rightarrow u=0.$$

则存在 $0\neq(x^*,y^*,z^*)$, 使得 $z^*\in N(0;\Theta)$ 满足关系 $(-x^*,-y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)$ 和

$$(x^*,y^*)\in\partial\langle u,q\rangle(\bar{x},\bar{y})+\bigcup_{\bar{v}\in M(\bar{x},\bar{y})}\Big[\nabla^2\langle\bar{v},g\rangle(\bar{x},\bar{y})(u)$$
$$+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\Big]$$

对某些 $u \in X \times Y$ 成立, 其中假设下述情形之一成立:

(a) f 在 $(\bar{x},\bar{y})$ 是 SNC 的, Θ 在 0 是 SNC 的;

(b) f 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, f^{-1} 在 $(\bar{z},\bar{x},\bar{y})$ 是强 PSNC 的, 这里, $\bar{z} := f(\bar{x},\bar{y})$.

更进一步, 当 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 则这些最优性条件等价于存在 $z^* \in N(0;\Theta) \setminus \{0\}$ 和 $u \in X \times Y$ 满足

$$0 \in \partial\langle z^*, f\rangle(\bar{x},\bar{y}) + \partial\langle u, q\rangle(\bar{x},\bar{y}) + \bigcup_{\bar{v}\in M(\bar{x},\bar{y})} \Big[\nabla^2\langle \bar{v}, g\rangle(\bar{x},\bar{y})(u)$$
$$+\nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{v})\big(\nabla g(\bar{x},\bar{y})u\big)\Big].$$

证明 在定理 5.81 中令 $Q(x,y) = \partial(\psi\circ g)(x,y)$, 通过对推论 3.76 所导出的顺从函数应用二阶次微分链式法则可得. △

定理 5.81 的最后一个推论关系到由具有复合域的参数广义方程控制的均衡/变分约束的 EPEC. 类似 5.2 节中的 MPEC, 这种约束可在最一般的情形进行研究. 简单起见, 下面只给出在某些光滑性假设下一类特殊的 EPEC 的必要最优性条件.

推论 5.84(具有复合域的 EPEC 的最优性条件) 设 $(\bar{x},\bar{y})$ 在约束

$$0 \in q(x,y) + (\partial\psi\circ g)(x,y)$$

下是局部 (f,Θ)- 最优的, 这里 $f\colon X\times Y \to Z$, $\Theta \subset Z$ 与前一个推论相同, 而 $g\colon X\times Y \to W$, $\psi\colon W \to \overline{\mathbb{R}}$, $q\colon X\times Y \to W^*$. 假设 X 和 Y 是 Asplund 空间, 同时 $\dim W < \infty$, q 和 g 在 $(\bar{x},\bar{y})$ 是严格可微的, 并且假设 $\mathrm{gph}\,\partial\psi$ 在 $(\bar{w},\bar{p})$ 附近是局部闭的, 其中 $\bar{w} = g(\bar{x},\bar{y})$, $\bar{p} = -q(\bar{x},\bar{y})$; 最后一个条件对连续或顺从函数来说自动成立. 还假设规范条件

$$\partial^2\psi(\bar{w},\bar{p})(0) \cap \ker\nabla g(\bar{x},\bar{y})^* = \{0\},$$

$$\Big[0 \in \nabla q(\bar{x},\bar{y})^*u + \nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{p})(u)\Big] \Rightarrow u = 0.$$

则在推论 5.83(a) 或 (b) 的情形下, 存在 $0 \neq (x^*,y^*,z^*)$ 且 $z^* \in N(0;\Theta)$, 满足

$$(x^*,y^*) \in D_N^*f(\bar{x},\bar{y})(z^*) \cap \Big[-\nabla q(\bar{x},\bar{y})^*u + \nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{p})(u)\Big]$$

对某个 $u \in X \times Y$ 成立.

更进一步, 当 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 则这些最优性条件等价于

$$0 \in \partial\langle z^*, f\rangle(\bar{x},\bar{y}) + \nabla q(\bar{x},\bar{y})^*u + \nabla g(\bar{x},\bar{y})^*\partial^2\psi(\bar{w},\bar{p})(u),$$

其中 $z^* \in N(0;\Theta) \setminus \{0\}$.

证明 这个结论在所给的假设下由当 $Q(x,y)=\partial(\psi\circ g)(x,y)$ 时定理 5.81 和定理 4.54 中所得的关于上导数 $D^*Q(\bar{x},\bar{y},\bar{p})$ 的上估计即得. △

正如 5.3.1 小节和 5.3.2 小节讨论的那样, 所得的结果直接蕴涵着具有特殊类型均衡的 EPEC 以及具有均衡约束的极小极大问题的必要最优性条件.

接下来将得到相对于闭序关系的 EPEC 的一些结果, 它们类似于但又独立于上面所给的结果. 和前面一样, 只给出所考虑的这些问题的点基/确切最优性条件. 先从具有由一般集值映射控制的抽象均衡约束的 EPEC 的最优性条件开始.

命题 5.85(具有闭序关系的抽象 EPEC 的最优性条件) 设 $(\bar{x},\bar{y})$ 是多目标最优化问题

$$\min f(x,y)\ (\text{相对于}\ \prec)\quad \text{s.t.}\quad y\in S(x)$$

的局部最优解, 这里 $f\colon X\times Y\to Z$ 是 Asplund 空间之间的映射, 其在 $(\bar{x},\bar{y})$ 附近是连续的且 $\bar{z}:=f(\bar{x},\bar{y})$, 偏好关系 $\prec$ 在 Z 上是闭的, 具有水平集 $\mathcal{L}(\cdot)$, $S\colon X\rightrightarrows Y$ 在 $(\bar{x},\bar{y})$ 附近是闭图的. 假设 f 在 $(\bar{x},\bar{y})$ 是 SNC 的, 或者 S 在该点是 SNC 的且 $\mathrm{cl}\mathcal{L}\colon Z\rightrightarrows Z$ 在 $(\bar{z},\bar{z})$ 是 ISNC 的. 则存在 $0\neq(x^*,y^*,z^*)\in X^*\times Y^*\times Z^*$ 满足

$$(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*),\quad -x^*\in D_N^*S(\bar{x},\bar{y})(y^*),\quad z^*\in N_+(\bar{z};\mathrm{cl}\mathcal{L}(\bar{z})).$$

更进一步, 如果 f 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 并且 $\dim Z<\infty$ 或者 S 在 $(\bar{x},\bar{y})$ 是 SNC 的且 $\mathrm{cl}\mathcal{L}$ 在 $(\bar{z},\bar{z})$ 是 ISNC 的, 则有

$$0\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+N\big((\bar{x},\bar{y});\mathrm{gph}\,S\big),\quad \text{其中}z^*\in N_+\big(\bar{z};\mathrm{cl}\mathcal{L}(\bar{z})\big)\setminus\{0\}.$$

证明 对约束集合是 Asplund 空间 $X\times Y$ 中的集合 $\Omega:=\mathrm{gph}S$ 的情形由定理 5.73(ii) 直接可得. △

现在已准备好得到涉及闭序关系和由参数变分系统/广义方程 (5.101) 控制的均衡约束的 EPEC 的必要最优性条件.

定理 5.86(具有闭序和变分约束的 EPEC 的最优性条件) 设 $(\bar{x},\bar{y})$ 是多目标最优化问题

$$\min f(x,y)\ (\text{相对于}\ \prec)\quad \text{s.t.}\quad 0\in q(x,y)+Q(x,y)$$

的局部最优解, 这里 $f\colon X\times Y\to Z$, $q\colon X\times X\to P$, $Q\colon X\times Y\rightrightarrows P$ 是 Asplund 空间之间的映射, $\prec$ 是 Z 上的闭序关系. 假设 f 和 q 在 $(\bar{x},\bar{y})$ 附近是连续的, Q 在 $(\bar{x},\bar{y},\bar{p})$ 附近是闭图的, 其中 $\bar{p}:=-q(\bar{x},\bar{y})\in Q(\bar{x},\bar{y})$. 如果下列情形之一满足:

(a) f 在 $(\bar{x},\bar{y})$ 是 SNC 的, Q 在 $(\bar{x},\bar{y},\bar{p})$ 是 SNC 的, 并且规范条件

$$\Big[(x^*,y^*)\in D_N^*g(\bar{x},\bar{y})(p^*)\cap\big(-D_N^*Q(\bar{x},\bar{y},\bar{p})(p^*)\big)\Big]\Rightarrow x^*=y^*=p^*=0$$

成立, 若 q 在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的, 它等价于 (5.102).

(b) f 在 $(\bar{x},\bar{y})$ 是 SNC 的, $\dim P<\infty$, q 在 $(\bar{x},\bar{y})$ 附近是 Lipschitz 连续的, 并且 (5.102) 满足.

(c) $\operatorname{cl}\mathcal{L}$ 在 $(\bar{z},\bar{z})$ 是 ISNC 的, g 在 $(\bar{x},\bar{y})$ 是 PSNC 的, 并且 (a) 中的规范条件成立.

则存在 $(x^*,y^*,z^*,p^*)\in X^*\times Y^*\times Z^*\times P^*$ 满足关系:

$$(x^*,y^*,z^*)\neq 0,\quad z^*\in N_+\big(\bar{z};\operatorname{cl}\mathcal{L}(\bar{z})\big),\quad \text{其中 } \bar{z}:=f(\bar{x},\bar{y}) \text{ 和}$$

$$(x^*,y^*)\in D_N^*f(\bar{x},\bar{y})(z^*)\cap\Big(-D_N^*q(\bar{x},\bar{y})(p^*)-D_N^*Q(\bar{x},\bar{y},\bar{p})(p^*)\Big).$$

更进一步, 对在 $(\bar{x},\bar{y})$ 是严格 Lipschitz 的 f 来说, 上面的最优性条件可等价地写为如下形式: 存在非零元 $z^*\in N_+\big(\bar{z};\operatorname{cl}\mathcal{L}(\bar{z})\big)$, 满足

$$0\in\partial\langle z^*,f\rangle(\bar{x},\bar{y})+D_N^*g(\bar{x},\bar{y})(p^*)+D_N^*Q(\bar{x},\bar{y},\bar{p})(p^*)$$

对某 $p^*\in P^*$ 成立. 此时 (a) 和 (b) 中 f 的 SNC 假设蕴涵着 $\dim Z<\infty$.

证明 当 S 由 (5.101) 给出时应用命题 5.85. 为此需根据 (5.101) 中的初始数据 (q,Q), 应用有效条件确保上导数 $D_N^*S(\bar{x},\bar{y})$ 的上估计和 S 在 $(\bar{x},\bar{y})$ 的 SNC 性质成立. 这可根据 4.4 节中的相应结果, 类似 5.81 的证明即可做到. △

类似上面广义序最优性的情形, 由定理 5.86 可得推论 5.82～ 推论 5.84 的相应结果, 给出涉及闭序关系和由上面所考虑的复合变分系统控制的均衡约束的 EPEC 的必要最优性条件.

5.4 线性率下的次极性和次优性

与以前考虑的问题相比, 本节主要研究具有更少限制的关于集合极性和标准最小化问题以及多目标最优化问题的 (次) 最优解的概念. 上面建立的对传统概念的必要极性和最优性条件对本节研究的新概念来说恰好是必要且充分的.

传统概念和下面将引入并研究的概念之间的主要差别在于, 后者关联于不是在所论点处而是在它的邻域内的极性/最优性, 以及它们涉及下面所给的精确定义意义下的线性率. 在某种程度上, 这类似于开性的线性率, 它把定义 1.51 中描述的覆盖性质和经典开映射定理框架下的一般开性区别开. 另外还要指出一般连续性和 Lipschitz 连续性之间的关系; 后者实际上意味着“具有线性率的连续性”. 正如覆盖和 Lipschitz 性质具有完全对偶刻画, 具有线性率的恰当定义的极性和最优性概念也有类似的刻画. 本节的主要目的是给出这些恰当的定义, 澄清它们的特殊性质, 并且证明相应的必要和充分极性/最优性条件.

本节首先从集合极性开始, 定义集合系统的线性次极性 (或线性率次极性), 并且阐明这样的系统可由极点原理的广义 Euler 方程以近似和确切的形式完全刻画. 接着考虑约束多目标最优化问题的线性次优性, 并且得到由上导数表示的关于这个概念的必要和充分条件. 本节的最后部分主要研究利用下半连续函数的次微分刻画它们的线性次极小性, 随后得到约束问题中的线性次极小性的必要和充分条件. 本节还用引人注目的例子说明了实值函数的标准极小性和线性次极小性之间的本质差别. 值得注意的是, 对严格可微函数来说, 线性次极小性约化为经典的驻点性质, 即严格导数在参考点为零.

5.4.1　集合系统的线性次极性

给定 Banach 空间 X 中的两个子集 Ω_1 和 Ω_2, 考虑常数

$$\vartheta(\Omega_1,\Omega_2):=\sup\big\{\nu\geqslant 0\mid \nu\mathbb{B}\subset\Omega_1-\Omega_2\big\},\tag{5.103}$$

它描述了这两个集合交叠性的度量. 注意到若 $\Omega_1\cap\Omega_2=\varnothing$, 则在 (5.103) 中有 $\vartheta(\Omega_1,\Omega_2)=-\infty$. 易见点 $\bar{x}\in\Omega_1\cap\Omega_2$ 是集合系统 $\{\Omega_1,\Omega_2\}$ 在定义 2.1 意义下的局部极点当且仅当

$$\vartheta\big(\Omega_1\cap B_r(\bar{x}),\Omega_2\cap B_r(\bar{x})\big)=0\tag{5.104}$$

对某 $r>0$ 成立, 其中照例 $B_r(\bar{x})=\bar{x}+r\mathbb{B}$. 修正 (5.104) 中的常数 $\vartheta(\cdot,\cdot)$, 则得到 Banach 空间中由两个集合构成的系统的“线性次极性”的概念如下.

定义 5.87(两个集合的线性次极性)　给定 $\Omega_1,\Omega_2\subset X$ 且 $\bar{x}\in\Omega_1\cap\Omega_2$. 称集合系统 $\{\Omega_1,\Omega_2\}$ 在 $\bar{x}$ 点附近是“线性次极值”的, 如果

$$\vartheta_{\rm lin}(\Omega_1,\Omega_2,\bar{x})=0,\tag{5.105}$$

其中

$$\vartheta_{\rm lin}(\Omega_1,\Omega_2,\bar{x}):=\liminf_{x_i\xrightarrow{\Omega_i}\bar{x},\,i=1,2,\,r\downarrow 0}\frac{\vartheta([\Omega_1-x_1]\cap r\mathbb{B},[\Omega_2-x_2]\cap r\mathbb{B})}{r},$$

其中交叠性的度量 $\vartheta(\cdot,\cdot)$ 由 (5.103) 定义.

显然在 (5.104) 意义下的集合极性蕴涵着 (5.105) 意义下的线性次极性, 但反之不然. 下面讨论集合系统线性次极性的一些特殊性质, 这区别于 (5.104) 中的概念:

(a) (5.105) 中定义的常数 $\vartheta_{\rm lin}(\Omega_1,\Omega_2,\bar{x})$ 与 (5.104) 中的常数 $\vartheta(\Omega_1\cap B_r(\bar{x}),\Omega_2\cap B_r(\bar{x}))$ 相比, 涉及当 $r\downarrow 0$ 时集合扰动的线性率. 因此条件 (5.105) 描述了集合 Ω_1 和 Ω_2 依线性率的局部非交叠性, 而 (5.104) 中的条件相当于这两个集合在 $r\downarrow 0$ 时以任意速率的局部非交叠性.

(b) 条件 (5.105) 要求的不是集合精确的局部非交叠性, 而是它们无限小变形的非交叠性. 这一点由下述表达式可得

$$\vartheta_{\text{lin}}(\Omega_1,\Omega_2,\bar{x}) = \liminf_{x_i \xrightarrow{\Omega_i} \bar{x},\, i=1,2} \vartheta_{\text{lin}}(\Omega_1 - x_1, \Omega_2 - x_2),$$

其中

$$\vartheta_{\text{lin}}(\Omega_1,\Omega_2) := \liminf_{r\downarrow 0} \frac{\vartheta(\Omega_1\cap r\mathbb{B}, \Omega_2\cap r\mathbb{B})}{r},$$

而 $\vartheta(\cdot,\cdot)$ 由 (5.103) 定义.

(c) 条件 (5.105) 不要求集合 Ω_1 和 Ω_2 在 $\bar{x}$ 点确切非交叠. 可以看到, 如果给定 $\bar{x}$ 的任意邻域 U, 存在点 $x_1 \in \Omega_1 \cap U$ 和 $x_2 \in \Omega_2 \cap U$ 保证平移集合 $\Omega_1 - x_1$ 和 $\Omega_2 - x_2$ 具有线性率的近似非叠加性, 则由 (b) 中的关系可知 (5.105) 成立.

在定理 2.20 中已经证明了对任意 Asplund 空间, 定义 2.5 中近似形式的极点原理的关系给出了定义 2.1 意义下局部集合极性的必要条件, 而定义 2.1 等价地由 (5.104) 描述. 实际上, 这些关系对上面定义的线性集合次极性来说恰好是必要和充分的. 确切的陈述在下面的定理中给出.

定理 5.88(利用近似极点原理刻画线性次极性)　设 Ω_1 和 Ω_2 是 Banach 空间 X 的子集, $\bar{x} \in \Omega_1 \cap \Omega_2$. 则下面的断言成立:

(i) 假设对任意正数 ε, 存在 $\widehat{x}_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ 和 $x_i^* \in \widehat{N}_\varepsilon(\widehat{x}_i;\Omega_i)\,(i=1,2)$, 满足

$$\|x_1^* + x_2^*\| \leqslant \varepsilon \quad 和 \quad \|x_1^*\| + \|x_2^*\| = 1, \tag{5.106}$$

则 $\{\Omega_1,\Omega_2\}$ 在 $\bar{x}$ 附近是线性次极值的.

(ii) 反之, 假设两集合 Ω_i 是局部闭的, 系统 $\{\Omega_1,\Omega_2\}$ 在 $\bar{x}$ 附近是线性次极值的. 如果 X 是 Asplund 空间, 则对任意 $\varepsilon > 0$, 存在 $\widehat{x}_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ 和 $x_i^* \in \widehat{N}_\varepsilon(\widehat{x}_i;\Omega_i)(i=1,2)$, 满足 (5.106). 而且, 如果这一性质对任意在某点 $\bar{x} \in \Omega_1 \cap \Omega_2$ 附近线性次极性系统 $\{\Omega_1,\Omega_2\} \subset X$ 来说都成立, 则 X 必为 Asplund 空间.

证明　为证 (i), 假设 $\{\Omega_1,\Omega_2\}$ 在 $\bar{x}$ 附近不是线性次极值的, 即对 (5.105) 中的常数 ϑ_{lin}, 有

$$\vartheta_{\text{lin}}(\Omega_1,\Omega_2,\bar{x}) =: \alpha > 0.$$

这意味着存在 $\bar{r} > 0$, 使得

$$\vartheta\big([\Omega_1 - x_1]\cap r\mathbb{B}, [\Omega_2 - x_2]\cap r\mathbb{B}\big) > (\alpha r)/2 \tag{5.107}$$

对任意正数 $r \leqslant \bar{r}$ 和任意 $x_i \in \Omega_i \cap r\mathbb{B}\quad (i=1,2)$ 成立, 其中 $\vartheta(\cdot,\cdot)$ 由 (5.103) 定义. 另一方面, 由当 $\varepsilon := \min\{\alpha/16, 1/4\}$ 时 (i) 中假设的条件和 (1.1) 中 ε- 法向量的定

义 (实际上它非常适合线性率的次极性的研究), 可得正数 $r < \bar{r}$ 满足

$$\langle x_i^*, x \rangle \leqslant \frac{\alpha}{32}\|x\|, \quad \forall x \in [\Omega_i - x_i] \cap r\mathbb{B}, \quad i = 1, 2.$$

由于 $x_2^* = -x_1^* + (x_1^* + x_2^*)$, 有

$$-\langle x_1^*, x \rangle \leqslant \left(\frac{\alpha}{32} + \varepsilon\right)\|x\| \leqslant \frac{3\alpha}{32}\|x\|, \quad \forall x \in [\Omega_2 - x_2] \cap r\mathbb{B} \text{ 和}$$

$$\langle x_1^*, x \rangle \leqslant (\alpha r)/8, \quad \forall x \in \big[(\Omega_1 - x_1) \cap r\mathbb{B}\big] - \big[(\Omega_2 - x_2) \cap r\mathbb{B}\big].$$

现由 (5.107) 和 (5.106) 中的第一个关系式得

$$\|x_1^*\| \leqslant 1/4, \quad \|x_2^*\| \leqslant \|x_1^*\| + \varepsilon \leqslant 1/2,$$

与 (5.106) 中的第二个关系式矛盾, 从而证明了断言 (i) 成立.

为证 (ii), 下面沿用引理 2.32(ii) 和定理 2.51(i) 的证明, 这与集合极性的必要条件相关. 相同的思想恰好对具线性率的集合次极性也适用.

设 $\{\Omega_1, \Omega_2\}$ 在 $\bar{x}$ 附近是线性次极值的, 即 (5.105) 成立. 给定 $\varepsilon \in (0,1)$, 找到 $x_i \in \Omega_i \cap (\varepsilon/2)\mathbb{B} \quad (i = 1, 2)$ 和 $0 < r < \varepsilon$, 使得

$$\vartheta\Big([\Omega_1 - x_1] \cap r\mathbb{B}, [\Omega_2 - x_2] \cap r\mathbb{B}\Big) < (r\varepsilon)/8.$$

根据交叠常数 $\vartheta(\cdot,\cdot)$ 的定义 (5.103), 这蕴涵着存在 $a \in (r\varepsilon/8)\mathbb{B}$ 满足

$$a \notin \big([\Omega_1 - x_1] \cap r\mathbb{B}\big) - \big([\Omega_2 - x_2] \cap r\mathbb{B}\big).$$

因此有

$$\|u - x_1 - v + x_2 - a\| > 0, \quad \text{如果 } u \in \Omega_1 \cap (x_1 + r\mathbb{B}), \quad v \in \Omega_2 \cap (x_2 + r\mathbb{B}).$$

由于 X 是 Asplund 空间, 则乘积空间 $X \times X$ 也是 Asplund 空间; 方便起见, 对它取最大模范数 $\|(u,v)\| := \max\{\|u\|, \|v\|\}$. 定义 $X \times X$ 上的一实值函数为

$$\varphi(u,v) := \|u - x_1 - v + x_2\|, \quad (u,v) \in X \times X,$$

则

$$\varphi(u,v) > 0, \quad \forall\, (u,v) \in \Omega := \Big[\Omega_1 \cap (x_1 + r\mathbb{B})\Big] \times \Big[\Omega_2 \cap (x_2 + r\mathbb{B})\Big],$$

其中 $\varphi(x_1, x_2) = \|a\| \leqslant (r\varepsilon)/8$.

由 Ekeland 变分原理 (定理 2.26) 知, 存在 $\bar{u} \in \Omega_1 \cap (x_1 + (r/4)\mathbb{B})$ 和 $\bar{v} \in \Omega_2 \cap (x_2 + (r/4)\mathbb{B})$, 使得 $(\bar{u},\ \bar{v})$ 是增广实值函数

$$\varphi(u,v) + \frac{\varepsilon}{2}\big\|(u,v) - (\bar{u},\bar{v})\big\| + \delta\big((u,v);\Omega\big), \quad (u,v) \in X \times X$$

的最小值点. 现在在任意 Asplund 空间中应用由引理 2.32(i) 给出的近似极点原理的次梯度描述, 并且考虑到上面和式中前两项都是凸的, 可找到

$$(y_1, y_2) \in (\bar{u}, \bar{v}) + \frac{r}{4}\mathbb{B} \subset (x_1, x_2) + \frac{r}{2}\mathbb{B}, \quad (z_1, z_2) \in \Omega \cap \big[(\bar{u}, \bar{v}) + \frac{r}{4}\mathbb{B}\big],$$

和 $(x_{1j}^*, x_{2j}^*), j = 1, 2, 3$, 满足

$$(x_{11}^*, x_{21}^*) \in \widehat{\partial}\varphi(y_1, y_2), \quad \|(x_{12}^*, x_{22}^*)\| \leqslant \varepsilon/2,$$

$$x_{13}^* \in \widehat{N}(z_1; \Omega_1), \quad x_{23}^* \in \widehat{N}(z_2; \Omega_2),$$

$$\big\|(x_{11}^*, x_{21}^*) + (x_{12}^*, x_{22}^*) + (x_{13}^*, x_{23}^*)\big\| \leqslant \varepsilon/2.$$

而且, $x_{11}^* = -x_{21}^* =: x^*$, 这里, x^* 是范数函数在非零点 $y_1 - x_1 - y_2 + x_2 - a$ 计算出的次梯度. 因此 $\|x^*\| = 1$, 并且

$$\|x_{13}^* + x_{23}^*\| \leqslant \|x_{13}^* + x^*\| + \|x_{23}^* - x^*\| = \big\|(x_{11}^*, x_{21}^*) + (x_{13}^*, x_{23}^*)\big\| \leqslant \varepsilon,$$

$$2 - \varepsilon \leqslant \|x_{13}^*\| + \|x_{23}^*\| \leqslant 2 + \varepsilon.$$

记 $\widehat{x}_1 := x_{13}$, $\widehat{x}_2 := x_{23}$,

$$x_1^* := x_{13}^*/\big(\|(x_{13}^*\| + \|x_{23}^*)\|\big), \quad x_2^* := x_{23}^*/\big(\|(x_{13}^*\| + \|x_{23}^*)\|\big).$$

则有 $x_i^* \in \widehat{N}(\widehat{x}_i; \Omega_i)$, $\widehat{x}_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ $(i = 1, 2)$, 且

$$\|x_1^*\| + \|x_2^*\| = 1, \quad \|x_1^* + x_2^*\| \leqslant \varepsilon/(2 - \varepsilon) \leqslant \varepsilon,$$

这就给出了 Asplund 空间中线性次极性系统近似极点原理的所有关系. 定理的最后一个陈述由定理 2.20 中 (b)$\Rightarrow$(a) 可得. △

下一个结果是定理 5.88 的推论, 在更多假设下利用确切极点原理的关系刻画了集合系统的线性次优性.

定理 5.89(利用确切极点原理刻画线性次极性) 设由 $\Omega_1, \Omega_2 \subset X$ 组成的系统在 $\bar{x} \in \Omega_1 \cap \Omega_2$ 附近是线性次极值的. 假设 X 是 Asplund 空间, 集合 Ω_1 和 Ω_2 在 $\bar{x}$ 附近是局部闭的, 并假设其中之一在该点是 SNC 的, 则存在 $x^* \in X^*$ 满足

$$0 \neq x^* \in N(\bar{x}; \Omega_1) \cap \big(-N(\bar{x}; \Omega_2)\big). \tag{5.108}$$

进一步, 如果 $\dim X < \infty$, 则条件 (5.108) 对 $\{\Omega_1, \Omega_2\}$ 在 $\bar{x}$ 附近的线性次极性来说是必要和充分的.

证明 首先根据定理 5.88(ii) 证明定理的第一个结论成立. 取 $\varepsilon_k \downarrow 0 \quad (k \to \infty)$, 利用定理 5.88(ii) 的结果, 找到序列 $x_{ik} \to \bar{x}$, $x_{ik}^* \in \widehat{N}_{\varepsilon_k}(x_{1k}; \Omega_i)$ $(i = 1, 2)$ 使得

$$\|x_{1k}^* + x_{2k}^*\| \leqslant \varepsilon_k, \quad \|x_{1k}^*\| + \|x_{2k}^*\| = 1, \quad \forall k \in \mathbb{N}. \tag{5.109}$$

由于 X 是 Asplund 空间, 序列 $\{x_{1k}^*\}$ 和 $\{x_{2k}^*\}$ 在 X^* 中有界, 则存在它们的子序列分别弱 * 收敛于 x_1^* 和 x_2^*. 由 (5.109) 中的第一个关系式和范数函数在 X^* 的弱 * 拓扑下的下半连续性得 $x_1^* = -x_2^* =: x^*$. 而且, 根据基本法锥的定义有 $x^* \in N(\bar{x}; \Omega_1) \cap (-N(\bar{x}; \Omega_2))$. 余下需证明如果两集合其中之一 (不妨设为 Ω_1) 在 $\bar{x}$ 是 SNC 的, 则 $x^* \neq 0$.

相反, 假设 $x^* = 0$. 则有 $x_{1k}^* \xrightarrow{w^*} 0$, 根据 Ω_1 在 $\bar{x}$ 的 SNC 性质, 有 $\|x_{1k}^*\| \to 0$. 由 (5.109) 中的第一个关系式也得 $\|x_{2k}^*\| \to 0$. 这显然与 (5.109) 中的第二个关系式矛盾, 从而完成了关于 Asplund 空间中闭集合的线性次极性系统的刻画 (5.108) 的证明.

现假设当 X 是有限维空间时 (5.108) 对 $\{\Omega_1, \Omega_2, \bar{x}\}$ $(\|x^*\| = 1)$ 成立. 利用有限维空间中基本法锥的表示 (1.8), 找到序列 $x_{ik} \xrightarrow{\Omega_i} \bar{x}$, $x_{1k}^* \to x^*$, $x_{2k}^* \to -x^*$, 使得 $x_{ik}^* \in \widehat{N}(x_{ik}; \Omega_i)$ 对 $i = 1, 2$ 和任意 $k \in \mathbb{N}$ 成立. 由于 $x_{1k}^* + x_{2k}^* \to 0$, $\|x_{1k}^*\| + \|x_{2k}^*\| \to 2\|x^*\| = 2 \quad (k \to \infty)$, 由标准正规化得对任意的 $\varepsilon > 0$, 存在 $x_i \in \Omega_i \cap (\bar{x} + \varepsilon\mathbb{B})$ 和 $x_i^* \in \widehat{N}(x_i; \Omega_i)$ $(i = 1, 2)$ 满足 (5.106). 因此根据定理 5.88(i), $\{\Omega_1, \Omega_2\}$ 在 $\bar{x}$ 附近是线性次极值的. 这就完成了定理的证明. △

值得注意的是, 定理 5.89 中第二部分的上述证明本质上利用了空间 X 的有限维性以确保 X^* 上的范数和弱 * 拓扑相同; 请比照 1.1.3 小节讨论的基础的 Josefson-Nissenzweig 定理. 另一方面, 该假设对具有特殊泛函结构的集合 Ω_i 来说可以放松; 更多细节参见下面两个小节.

注 5.90(多个集合的线性次极性) 线性集合次极性的上述定义涉及两个集合的情形. 在 Banach 空间 X 中给定包含有限多个集合的系统 $\{\Omega_1, \cdots, \Omega_n\}$ $(n \geqslant 2)$, 可用下列方式定义线性次极性: $\{\Omega_1, \cdots, \Omega_n\}$ 在 $\bar{x} \in \Omega_1 \cap \cdots \cap \Omega_n$ 附近是线性次极值的, 如果包含如下两个集合

$$\widetilde{\Omega}_1 := \prod_{i=1}^{n} \Omega_i \quad \text{和} \quad \widetilde{\Omega}_2 := \{(x, \cdots, x) \in X^n \mid x \in X\}$$

的系统在定义 5.87 的意义下在 $(\bar{x}, \cdots, \bar{x}) \in X^n$ 附近是线性次极值的. 这可以等价地描述为, 给定任意的 $j \in \{1, \cdots, n\}$, 两个集合

$$\overline{\Omega}_1 := \prod_{i \in \{1, \cdots, n\} \setminus \{j\}} \Omega_i \quad \text{和} \quad \overline{\Omega}_2 := \big\{(x, \cdots, x) \in X^{n-1} \mid x \in \Omega_j\big\}$$

的系统在 $(\bar{x},\cdots,\bar{x})\in X^{n-1}$ 附近是线性次极值的.

根据两个集合情形的上述结果和初等的计算, 可得定理 5.88 和定理 5.89 关于有限多个集合系统的相应结果. 特别地, Asplund 空间 X 中的局部闭集系统 $\{\Omega_1,\cdots,\Omega_n\}$ $(n\geqslant 2)$ 在 $\bar{x}$ 附近是线性次极值的当且仅当近似极点原理的下面关系成立: 对任意的 $\varepsilon>0$, 存在 $x_i\in\Omega_i\cap(\bar{x}+\varepsilon\mathbb{B})$ 和 $\widehat{N}_\varepsilon(x_i;\Omega_i)$ $(i=1,\cdots,n)$ 满足

$$\|x_1^*+\cdots+x_n^*\|\leqslant\varepsilon,\quad \|x_1^*\|+\cdots+\|x_n^*\|=1.$$

而且, 如果除一个外所有的 Ω_i 在 $\bar{x}$ 是 SNC 的, 则对在 $\bar{x}$ 附近是线性次极值的任何系统 $\{\Omega_1,\cdots,\Omega_n\}$, 有确切极点原理的关系:

存在 $x_i^*\in N(\bar{x};\Omega_i)$ $(i=1,\cdots,n)$ 满足

$$x_1^*+\cdots+x_n^*=0,\quad \|x_1^*\|+\cdots+\|x_n^*\|=1.$$

更进一步, 当 X 是有限维的, 后面的关系式对 $\{\Omega_1,\cdots,\Omega_n\}$ 在 $\bar{x}$ 附近的线性次极性来说是必要和充分的.

5.4.2 多目标最优化中的线性次优性

本小节考虑约束多目标最优化的一些问题并且研究这些问题的线性次优解的新概念. 这个概念与前一小节研究的集合系统的线性次极性 (类似 5.3.1 小节中广义序最优性和集合极性之间的关系) 密切相关 (实际上由其诱导), 尽管这里利用原始数据独立地描述它. 本小节基本目标是得到线性次优解的近似/模糊和确切/点基两种形式的必要和充分条件 (以及仅仅必要的条件). 尽管前一种形式的条件将在更一般的假设下得到, 由于具有应用充分发展的关于基本法锥/上导数/次微分结构的分析法则的可能性, 后一条件有更多优势. 这对研究各种约束的多目标问题来说至关重要.

给定 Banach 空间之间的映射 $f\colon X\to Z$, 子集 $\Omega\subset X$ 和 $\Theta\subset Z$, 点 $\bar{x}\in\Omega$ 且 $f(\bar{x})\in\Theta$, 引入常数

$$\vartheta_{\rm lin}(f,\Omega,\Theta,\bar{x}):=\liminf_{x\xrightarrow{\Omega}\bar{x},\,z\xrightarrow{\Theta}f(x),\,r\downarrow 0}\frac{\vartheta\Big(f\big(B_r(x)\cap\Omega\big)-f(x),\Theta-z\Big)}{r},\tag{5.110}$$

这里 $\vartheta(\cdot,\cdot)$ 由 (5.103) 定义.

定义 5.91(多目标问题的线性次优解) 给定 $(f,\Omega,\Theta,\bar{x})$ 如上, 称 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解, 如果有

$$\vartheta_{\rm lin}(f,\Omega,\Theta,\bar{x})=0,$$

其中常数 $\vartheta_{\rm lin}(f,\Omega,\Theta,\bar{x})$ 由 (5.110) 定义.

易验证, 在定义 5.53 意义下 (简单起见, 令 $f(\bar{x})=0$), 每个约束 $x\in\Omega$ 的局部 (f,Θ)- 最优的 $\bar{x}$ 恰好也相对 (f,Ω,Θ) 线性次优. 因此线性次优解的上述概念是 5.3.5 小节中研究的约束多目标问题的 (确切) 广义序最优性的推广. 除了次优性和最优性的区别, 定义 5.91 和定理 5.53 中解概念之间的另一至关重要的差别在于线性率; 请比照定义 5.87 后面关于集合极性和线性次极性之间关系的讨论. 这允许在一般情形下得到线性次优解的必要和充分条件. 下面首先得到该方向的一个“模糊”结果, 这个结果与定理 5.88 中的近似极点原理的线性次极性刻画密切相关 (实际上等价). 为描述这个结果, 定义关于 (f,Ω,Θ) 的集值映射 $F\colon X\rightrightarrows Z$,

$$F(x):=\begin{cases} f(x)-\Theta, & \text{若 } x\in\Omega, \\ \varnothing, & \text{其他}. \end{cases} \tag{5.111}$$

值得注意的是, 这个映射 F 的图与 5.3.2 小节考虑的广义上图集合 $\mathcal{E}(f,\Omega,\Theta)$ 相同.

定理 5.92(多目标最优化线性次优性的模糊刻画) 设 X 和 Y 是 Banach 空间, $\bar{x}\in\Omega$ 且 $f(\bar{x})\in\Theta$. 则下面的结论成立:

(i) 假设对任意的 $\varepsilon>0$, 存在 $(x,z)\in(\bar{x},0)+\varepsilon\mathbb{B}_{X\times Z}$, 其中 $z\in F(x)$, 存在 $z^*\in Z^*, 1-\varepsilon\leqslant\|z^*\|\leqslant 1+\varepsilon$ 满足包含关系

$$0\in\widehat{D}^*_\varepsilon F(x,z)(z^*). \tag{5.112}$$

则 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解.

(ii) 反之, 假设 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解. 如果 $\operatorname{gph}F$ 在 $(\bar{x},0)$ 附近是局部闭的, 空间 X 和 Z 都是 Asplund 空间, 则对任意的 $\varepsilon>0$, 存在 $(x,z)\in(\bar{x},0)+\varepsilon\mathbb{B}_{X\times Z}$ 且 $z\in F(x)$, $z^*\in Z^*$, $1-\varepsilon\leqslant\|z^*\|\leqslant 1+\varepsilon$ 满足包含关系

$$0\in\widehat{D}^*F(x,z)(z^*).$$

证明 易见 $\bar{x}$ 是对于 (f,Ω,Θ) 的线性次优解当且仅当包含两个集合

$$\Omega_1:=\operatorname{gph}F,\quad \Omega_2:=X\times\{0\}\subset X\times Z$$

的系统在 $(\bar{x},0)\in X\times Z$ 附近是线性次极值的. 对这个集合系统 $\{\Omega_1,\Omega_2\}$ 应用定理 5.88 中的线性次极性的刻画, 并且考虑到 $\widehat{N}_\varepsilon((x,0);\Omega_2)=(\varepsilon\mathbb{B}^*)\times Z^*$ 及

$$(0,-z^*)\in\widehat{N}_\varepsilon\big((x,z);\Omega_1\big)\Longleftrightarrow 0\in\widehat{D}^*_\varepsilon F(x,z)(z^*)$$

对任意 $\varepsilon\geqslant 0$ 成立, 即得到定理的所有结论. △

推论 5.93(线性次优性的模糊刻画的推论) 条件 (5.112) 总是蕴涵着 $z^* \in \widehat{N}_\varepsilon\big(f(x);\Theta\big)$ 对任意的 $\varepsilon \geqslant 0$ 成立, 而且, 对任意接近于 $\bar{x}$ 且满足 $z = f(x) \in \Theta$ 的 $x \in X$, 如果 f 在 $\bar{x}$ 附近相对于约束集合 Ω 是 Lipschitz 连续的, 则有

$$0 \in \widehat{D}^* F(x,z)(z^*) \Leftrightarrow 0 \in \widehat{\partial}\langle z^*, f_\Omega\rangle(x), \quad z^* \in \widehat{N}\big(f(x);\Theta\big),$$

其中 $f_\Omega = f + \delta(\cdot;\Omega)$.

证明 由局部 Lipschitz 函数的 Fréchet 上导数的定义和 (简单的) 标量化公式直接可得. △

下一定理以压缩形式, 即利用基于初始数据 (f,Ω,Θ) 在 (5.111) 中的映射 F 给出了多目标最优化问题的线性次优解的必要条件和充分条件 (以及点基刻画). 这些结果利用仅在参考解处 (而不是附近) 映射 F 的混合上导数 (1.25) 和反向混合上导数 (1.40) 来表示. 注意到加在定理 5.94(ii) 中映射 F^{-1} 的 PSNC 性质与定理 5.59 中集合 $\mathcal{E}(f,\Omega,\Theta)$ 的 PSNC 性质相同. 因此, 当 $\bar{z}=0$ 时, 定理 5.59(ii) 中假设 (a) 与 (b) 的任意其中之一保证 F^{-1} 在 $(0,\bar{x})$ 所要求的 PSNC 性质成立; 参见定理 5.59 的证明. 另外定理 5.94(iii) 中 F 的强上导数正规性的充分条件已罗列在命题 4.9 中.

定理 5.94(多目标问题线性次优性的压缩点基条件) 设 F 是基于 (f,Ω,Θ) 由 (5.111) 定义的 Banach 空间之间的映射. 则下列结论成立:

(i) 假设 $\dim X < \infty$, 且存在 $0 \neq z^* \in Z^*$ 满足

$$0 \in D^*_M F(\bar{x},0)(z^*),$$

则 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解.

(ii) 反之, 假设 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解. 如果 X 和 Z 都是 Asplund 空间, $\operatorname{gph} F$ 在 $(\bar{x},0)$ 附近是局部闭的, F^{-1} 在 $(0,\bar{x})$ 是 PSNC 的 (该条件自动成立若 $\dim Z < \infty$). 则存在 $0 \neq z^* \in Z^*$ 满足

$$0 \in \widetilde{D}^*_M F(\bar{x},0)(z^*).$$

(iii) 设 $\dim X < \infty$, Z 是 Asplund 空间, F 在 $(\bar{x},0)$ 附近是闭图的, 还假设 F 在 $(\bar{x},0)$ 是 SNC 的和强上导数正规的, 其中 $D^*F(\bar{x},0) := D^*_M F(\bar{x},0) = D^*_N F(\bar{x},0)$. 则 $\bar{x}$ 相对于 (f,Ω,Θ) 是线性次优解当且仅当存在 $0 \neq z^* \in Z^*$ 满足

$$0 \in D^* F(\bar{x},0)(z^*).$$

证明 先证 (i). 应用 $0 \in D^*_M F(\bar{x},0)(z^*)$ 和当 $\dim X < \infty$ 时混合上导数的定义, 找到 $\varepsilon_k \downarrow 0$, $x_k \to \bar{x}$, $z_k \to 0$, $x_k^* \to 0$ 和 $z_k^* \to z^*$, 使得

$$z_k \in F(x_k), \quad x_k^* \in \widehat{D}^*_{\varepsilon_k} F(x_k,z_k)(z_k^*), \quad \forall k \in \mathbb{N}.$$

注意到根据 (5.111) 中 F 的结构, 上面第一个包含关系蕴涵着 $x_k \in \Omega$, $z_k = f(x_k) \in \Theta$. 而且, 由于 $\|z_k^* - z^*\| \to 0$, $\|z^*\| = 1$, 不失一般性, 可假设 $\|z_k^*\| = 1$ $(\forall k \in \mathbb{N})$. 由 $x_k^* \in \widehat{D}_{\varepsilon_k}^* F(x_k, z_k)(z_k^*)$ 得

$$\langle x_k^*, x - x_k\rangle - \langle z_k^*, z - z_k\rangle \leqslant \varepsilon_k(\|x - x_k\| + \|z - z_k\|)$$

对任意充分接近于 $(\bar{x}, 0)$ 的 (x, z) 成立, 这蕴涵着估计

$$-\langle z_k^*, z - z_k\rangle \leqslant \big(\varepsilon_k + \|x_k^*\|\big)\big(\|x - x_k\| + \|z - z_k\|\big),$$

从而

$$0 \in \widehat{D}_{\gamma_k}^* F(x_k, z_k)(z_k^*), \quad 其中 \ \gamma_k := \varepsilon_k + \|x_k^*\| \downarrow 0 \quad (k \to \infty).$$

现应用定理 5.92(i), 得 $\bar{x}$ 相对于 (f, Ω, Θ) 是线性次优解.

为证 (ii), 取相对于 (f, Ω, Θ) 的线性次优解 $\bar{x}$, 并任取序列 $\varepsilon_k \downarrow 0$ $(k \to \infty)$. 利用定理 5.92(ii), 找到序列 $(x_k, z_k) \to (\bar{x}, 0)$ 且 $z_k \in F(x_k)$ 和 $z_k^* \in Z^*$ 且 $\|z_k^*\| = 1$ 满足 $0 \in \widehat{D}^* F(x_k, z_k)(z_k^*)$ $(\forall k \in \mathbb{N})$. 由于 Z 是 Asplund 空间, 存在 $z^* \in Z^*$ 使得 $z_k^* \xrightarrow{w^*} z^*(k \to \infty$, 沿子序列), 而且通过取极限显然有 $0 \in \widetilde{D}_M^* F(\bar{x}, 0)(z^*)$. 更进一步, 根据所给的 PSNC 假设有 $z^* \neq 0$. 如果 Z 是有限维的, 则 PSNC 的假设显然成立. 当 F 在 $(\bar{x}, 0)$ 附近是度量正则的时, 根据命题 1.68 及 F 的类 Lipschitz 性质和 F^{-1} 的度量正则性是等价的, 该假设也满足. 因此得 (ii) 中的所有结论.

最后的断言 (iii) 是 (i) 与 (ii) 的直接组合. 注意到由于 $\dim X$ 是有限维的, 此时 $\widetilde{D}_M^* F(\bar{x}, 0) = D_N^* F(\bar{x}, 0)$, F^{-1} 的 PSNC 性质等价与 F 的 SNC 性质. △

应用完全分析法则, 由定理 5.94(ii) 中的压缩结果, 可推导出分别通过初始数据 (f, Ω, Θ), 即利用 f, Ω 和 Θ 的广义微分结构, 来表示的在各种 (特别地, 均衡) 约束下的多目标问题及其特殊情形的线性次优性的必要条件; 请比照 5.3.2 小节和 5.3.5 小节中关于广义序最优性的结果. 充分条件以及线性次优性的刻画的情形是更精细的: 此时必须利用分析法则的等式形式, 它们本质上比对必要性的要求有更多的限制. 基于定理 5.94(iii) 的压缩条件, 下面利用初始数据 (f, Ω, Θ) 给出关于线性次优性的刻画的一些结果.

定理 5.95(多目标问题中线性次优性的分离型/非压缩型点基准则) 设 $f: X \to Z$ 在 $\bar{x}$ 附近是 Lipschitz 连续的且 $\dim X < \infty$, 设 $\Omega \subset X$ 和 $\Theta \subset Z$ 分别在 $\bar{x}$ 和 $\bar{z} := f(\bar{x}) \in \Theta$ 附近是局部闭的. 设加在初始数据上的下列假设 (a)~(c) 之一成立:

(a) $\dim Z < \infty$, 并且 $\Omega = X$ 或 f 在 $\bar{x}$ 是严格可微的.

(b) Z 是 Asplund 空间, $\Omega = X, \Theta$ 在 $\bar{z}$ 是法向量正则的和 SNC 的, 并且 f 在 $\bar{x}$ 是严格 Lipschitz 的.

(c) Z 是 Asplund 空间, Ω 在 $\bar{x}$ 是法向正则的, Θ 在 $\bar{z}$ 是法向正则的和 SNC 的, 并且 f 在 $\bar{x}$ 是 N- 正则的.

则 $\bar{x}$ 是相对于 (f,Ω,Θ) 的线性次优解当且仅当存在 $0\neq z^*\in Z^*$ 满足

$$0\in\partial\langle z^*,f\rangle(\bar{x})+N(\bar{x};\Omega),\quad z^*\in N(\bar{z};\Theta).$$

证明 由于 $\text{gph}F=\mathcal{E}(f,\Omega,\Theta)$, 其中 $\mathcal{E}(f,\Omega,\Theta)$ 由 (5.37) 定义, 则有

$$D_N^*F(\bar{x},0)(z^*)=\big\{x^*\in X^*\,\big|\,(x^*,-z^*)\in N\big((\bar{x},0);\mathcal{E}(f,\Omega,\Theta)\big)\big\}.$$

如果 Z 是 Asplund 空间, f 在 Ω 上的限制 f_Ω 在 $\bar{x}$ 附近是局部 Lipschitz 的, 并且在该点是强上导数正规的, 则由引理 5.23(iii) 和 (iv) 可得表达式

$$D_N^*F(\bar{x},0)(z^*)=\begin{cases}\partial\langle z^*,f_\Omega\rangle(\bar{x}), & \text{若} z^*\in N(\bar{z};\Theta),\\ \varnothing, & \text{其他}.\end{cases}\tag{5.113}$$

首先考虑 $\Omega=X$ 的情形. 由 (5.113) 可知, F 在 $(\bar{x},0)$ 是强上导数正规的, 如果 $\dim Z<\infty$, 或 f 在 $\bar{x}$ 是严格 Lipschitz 的且 Θ 在 $\bar{z}$ 是法向正则的; 参见命题 4.9. 为满足定理 5.94(iii) 中所有的假设, 还需验证 (在 $\dim Z=\infty$ 的情况下)F^{-1} 在 $(0,\bar{x})$ 是 PSNC 的. 利用定理 5.59(ii) 的证明, 如果 Θ 在 $\bar{z}$ 是 SNC 的, 或 f^{-1} 在 $(\bar{z},\bar{x})$ 是 PSNC 的, 则能保证这个性质成立. 由于 X 是有限维的, 后面的条件等价于 f 在 $(\bar{x},\bar{z})$ 的 SNC 性质, 根据推论 3.30, 对严格 Lipschitz 映射而言, 这个性质简化为 $\dim Z<\infty$. 这就在 $\Omega=X$ 的情形下完成了定理的证明.

对在 $\Omega\neq X$ 的约束情形, 需证 (5.113) 中的等式

$$\partial\langle z^*,f_\Omega\rangle(\bar{x})=\partial\langle z^*,f\rangle(\bar{x})+N(\bar{x};\Omega)$$

成立. 利用命题 1.107(ii) 的和法则, 当 f 在 $\bar{x}$ 严格可微时等式成立. 而且, 如果 f 在 $\bar{x}$ 是 N- 正则的, Ω 在该点是法向正则的, 则此等式成立, 且 f_Ω 在该点也是 N- 正则的 (从而是强上导数正规的); 参见命题 3.12 和 4.9. 这些事实与 $\dim Z=\infty$ 情形下所需的 (c) 中关于 Θ 的假设合并在一起, 类似 $\Omega=X$ 的上面的证明, 即得到定理 5.94(iii) 中的所有要求, 从而完成了定理的证明. △

下面建立上一定理的推论, 它给出具有算子约束的多目标问题的线性次优解的一个刻画. 值得注意的是, 在由集值和非光滑映射给出一般算子约束条件下, 推论 5.60 中所得的相应必要最优性条件不需任何改动即对线性次优解成立. 然而, 下面给出的必要和充分条件本质上要求对初始数据有更多的限制假设来保证关于逆像的分析法则中的等式成立, 而且在无限维情形保证逆像的法向正则性成立. 这不

可避免地将这里的结果限制于以严格可微映射来描述多目标问题中的算子和泛函约束.

推论 5.96(算子约束下线性次优性的点基准则) 设 $f\colon X \to Z$, $g\colon X \to Y$, $\Theta \subset Z$, $\Lambda \subset Y$. 假设 $\dim X < \infty$, Θ 和 Λ 分别在 $\bar{z}$ 和 $\bar{y} := g(\bar{x})$ 附近是局部闭的, f 和 g 在 $\bar{z}$ 是严格可微的. 还假设下列条件之一也成立:

(a) Y 是 Banach 空间, $\dim Z < \infty$, 并且 $\nabla g(\bar{x})$ 是满射.

(b) $\dim Y < \infty$, Z 是 Asplund 空间, Λ 在 $\bar{y}$ 是法向正则的, Θ 在 $\bar{z}$ 是法向正则且 SNC, 并且

$$N(\bar{y}; \Lambda) \cap \ker \nabla g(\bar{x})^* = \{0\}.$$

则 $\bar{x}$ 相对于 $(f, g^{-1}(\Lambda), \Theta)$ 是线性次优的当且仅当存在 $0 \neq z^* \in Z^*$ 满足

$$0 \in \nabla f(\bar{x})^* z^* + \nabla g(\bar{x})^* N(\bar{y}; \Lambda), \quad z^* \in N(\bar{z}; \Theta).$$

证明 当 $\Omega := g^{-1}(\Lambda)$ 时应用定理 5.95. 首先应用定理 1.17, 当 Y 是 Banach 空间, 在 (a) 中所给的关于 $\nabla g(\bar{x})$ 满射性假设下有分析法则公式

$$N(\bar{x}; \Omega) = \nabla g(\bar{x})^* N(\bar{y}; \Lambda)$$

成立. 然后由定理 5.95(a) 得此推论的结论成立.

除上面的分析法则外, 定理 5.95(c) 还需要 $\Omega = g^{-1}(\Lambda)$ 的法向正则性. 为此利用定理 3.13(iii), 其中 $F(y) = \delta(y; \Lambda)$. 这就证明了在 (b) 中所给假设下推论的结论成立. 值得注意的是, 由定理 3.13 中关于 g 的 N- 正则性条件只能得到 g 的严格可微性, 因为由 $\dim X < \infty$ 时的推论 3.69 知, g 的图正则性等价于它在参考点处的严格可微性. △

所得结果对具有由严格可微函数给出的等式和不等式经典形式的泛函约束的多目标问题情形有引人注目的推论. 此时标准/规范形式的 Lagrange 乘子法则的合适的多目标版本在 Mangasarian-Fromovitz 约束规范下给出了线性次优性的充要条件.

推论 5.97(具有泛函约束的多目标问题的线性次优性) 设 $f\colon X \to Z$ 在 $\bar{x}$ 是严格可微的且 $\dim X < \infty$, Z 是 Asplund 空间, Θ 在 $\bar{z}$ 是法向正则且 SNC 的, 并令

$$\Omega = \left\{ x \in X \,\middle|\, \varphi_i(x) \leqslant 0, \quad i = 1, \cdots, m; \quad \varphi_i(x) = 0, \quad i = m+1, \cdots, m+r \right\},$$

其中每个 $\varphi_i(x)$ 在 $\bar{x}$ 是严格可微的. 假设 Mangasarian-Fromovitz 约束规范成立:

(a) $\nabla \varphi_{m+1}(\bar{x}), \cdots, \nabla \varphi_{m+r}(\bar{x})$ 线性无关, 且

(b) 存在 $u \in X$ 满足

$$\langle \nabla \varphi_i(\bar{x}), u \rangle < 0, \quad i \in \{1, \cdots, m\} \cap I(\bar{x});$$

$$\langle \nabla \varphi_i(\bar{x}), u \rangle = 0, \quad i = m+1, \cdots, m+r,$$

其中 $I(\bar{x}) := \{i = 1, \cdots, m+r | \varphi_i(\bar{x}) = 0\}$.

则 $\bar{x}$ 相对于 (f, Ω, Θ) 是线性次优的当且仅当存在 $z^* \in N(\bar{z}; \Theta) \backslash \{0\}$, $(\lambda_1, \cdots, \lambda_{m+r}) \in \mathbb{R}^{m+r}$, 使得

$$\nabla f(\bar{x})^* z^* + \sum_{i=1}^{m+r} \lambda_i \nabla \varphi_i(\bar{x}) = 0,$$

$$\lambda_i \geqslant 0, \quad \lambda_i \varphi_i(\bar{x}) = 0 \quad (\forall i = 1, \cdots, m).$$

证明　置

$$\Lambda := \Big\{ (\alpha_1, \cdots, \alpha_{m+r}) \in \mathbb{R}^{m+r} \ \Big| \ \alpha_i \leqslant 0,\ i = 1, \cdots, m; \quad \alpha_i = 0,\ i = m+1, \cdots, m+r \Big\},$$

$$g := (\varphi_1, \cdots, \varphi_{m+r}) \colon X \to \mathbb{R}^{m+r}$$

并应用推论 5.96(b).　△

下面建立具有均衡约束的多目标问题, 即 EPEC(在 5.3.5 小节中的术语下) 的线性次优性的充要条件. 考虑到上述讨论, 这样问题的一般框架描述如下：给定 $f \colon X \times Y \to Z$, $S \colon X \rightrightarrows Y$, $\Theta \subset Z$, 称 $(\bar{x}, \bar{y})$ 相对于 (f, Ω, Θ) 是线性次优的, 如果它在定义 5.91 意义下相对于 $(f, \operatorname{gph} S, \Theta)$ 是线性次优的. 这里主要关注由形为 $0 \in q(x, y) + Q(x, y)$ 的参数变分系统的解映射描述的均衡约束.

首先注意到, 根据定理 5.94(ii) 和包含型的分析法则, 5.3.5 小节中所得的关于广义序最优性的所有必要条件对所考虑的 EPEC 的线性次优解来说也成立. 为得到线性次优性准则, 则需利用更多限制的等式型分析法则, 它们给出了计算由均衡约束给出的解映射的上导数的确切公式, 而且还保证在合适情形下这些映射的图正则性. 为此, 利用当 $\Omega = \operatorname{gph} S \subset X \times Y$ 时定理 5.95 中的结果, 及在 4.4.1 小节建立的相应的上导数公式和关于参数变分系统的正则性断言. 为简单起见, 在下一定理中对 f 要求严格可微性 (替代定理 5.95(c) 中的 N- 正则性), 当 $\dim Z < \infty$ 时, 这个条件是不可避免的, 而在无限维中它可以被放宽, 参看定理 3.68.

定理 5.98(一般 EPEC 线性次优性的刻画)　设 $f \colon X \times Y \to Z$, $q \colon X \times Y \to P$ 在 $(\bar{x}, \bar{y})$ 是严格可微的, 且 $\bar{z} := f(\bar{x}, \bar{y}) \in \Theta$, $\bar{p} := -q(\bar{x}, \bar{y})$; 设 $\Theta \subset Z$ 和 $Q \colon X \times Y \rightrightarrows$

P 的图分别在 $\bar{z}$ 和 $(\bar{x},\bar{y},\bar{p})\in \mathrm{gph}\, Q$ 附近是局部闭的; 设 X 和 Y 都是有限维的; 并且设

$$S(x):=\big\{y\in Y \,\big|\, 0\in q(x,y)+Q(x,y)\big\}.$$

此外, 假设下列条件之一成立:

(a) $\dim Z<\infty$, P 是 Banach 空间, $\nabla_x q(\bar{x},\bar{y})$ 是满射, 并且 $Q=Q(y)$.

(b) Z 和 P 是 Asplund 空间, Θ 在 $\bar{z}$ 是 SNC 和法向正则的, $Q=Q(x,y)$ 在 $(\bar{x},\bar{y},\bar{p})$ 是 SNC 和 N- 正则的, 而且伴随广义方程

$$0\in \nabla q(\bar{x},\bar{y})^* p^* + D_N^* Q(\bar{x},\bar{y},\bar{p})(p^*)$$

只有平凡解 $p^*=0$.

则 $(\bar{x},\bar{y})$ 相对于 (f,S,Θ) 是线性次优的当且仅当存在 $z^*\in N(\bar{z};\Theta)\backslash\{0\}$, $p^*\in P^*$ 满足

$$0\in \nabla f(\bar{x},\bar{y})^* z^* + \nabla q(\bar{x},\bar{y})^* p^* + D_N^* Q(\bar{x},\bar{y},\bar{p})(p^*).$$

证明 当 $\Omega=\mathrm{gph}\, S\subset X\times Y$ 时利用定理 5.95. 如果 X 和 Y 都是有限维的, f 在 $(\bar{x},\bar{y})$ 是严格可微的, $\dim Z<\infty$ 或 Z 是 Asplund 空间, 且 Θ 在 $\bar{z}$ 是 SNC 和法向正则的, 及 S 在 $(\bar{x},\bar{y})$ 是 N- 正则的, 则 $(\bar{x},\bar{y})$ 相对于 (f,S,Θ) 是线性次优的当且仅当存在 $z^*\in N(\bar{z};\Theta)\backslash\{0\}$ 满足

$$0\in \nabla f(\bar{x},\bar{y})^* z^* + N\big((\bar{x},\bar{y});\mathrm{gph}\, S\big).$$

为通过初始数据得到关于解映射 S 的结果, 需要利用 (q,Q) 来表示 $N\big((\bar{x},\bar{y});\mathrm{gph}\, S\big)$; 当 $\dim Z=\infty$ 时, 还需利用额外条件使得 S 在 $(\bar{x},\bar{y})$ 的 N- 正则性成立. 首先考虑 $\dim Z<\infty$ 的情况, 此时不需要 S 的正则性. 在这种情形下, 当 P 是 Banach 空间, $Q=Q(\bar{y})$, 且 $\nabla_x q(\bar{x},\bar{y})$ 是满射的时, 根据定理 4.44(i) 得

$$N\big((\bar{x},\bar{y});\mathrm{gph}\, S\big)=\big\{(x^*,y^*)\in X^*\times Y^* \,\big|\, x^*=\nabla_x q(\bar{x},\bar{y})^* p^*,$$
$$y^*\in \nabla_y q(\bar{x},\bar{y})^* p^* + D_N^* Q(\bar{y},\bar{p})(p^*)\ \ \text{对某个}p^*\in P^*\text{成立}\big\}.$$

这就给出了情形 (a) 下定理的结论.

如果 $Q=Q(x,y)$, Z 是 Asplund 空间, 利用定理 4.44(ii), 在关于 Q 的正则性假设但 $\nabla_x q(\bar{x},\bar{y})$ 不具满射性条件下, 它给出了 $N((\bar{x},\bar{y});\mathrm{gph}\, S)$ 的表示公式, 而且同时确保 S 在 $(\bar{x},\bar{y})$ 的 N- 正则性成立. 合并这个结论与定理 5.95(c) 中的假设, 就完成了定理的证明. △

定理 5.98(b) 中最有限制性的假设是域 Q 的 N- 正则性. 特别地, 当 Q 是凸图的, 该性质成立. 在这种情况下由推论 4.45 易得定理 5.98 的特例, 其中的描述用到 Q, 而不是其上导数.

下面给出当 $Q = \partial(\psi \circ g)$ 时定理 5.98 的一个特例, 即当所考虑的广义方程的域具有复合势且以次微分形式给出. 正如 4.4.1 小节中讨论的那样, 这样的模型包含经典的变分不等式及推广. 为得到这种类型的 EPEC 的线性次优性的刻画, 需要二阶次微分链式法则以便通过初始数据 (ψ, g) 来表示 $D^*Q = \partial^2(\psi \circ g)$. 和前述一样, 只可用那些能保证等式型链式法则的分析法则. 由于图正则性对具有非光滑势的次可微映射来说不是实际可行的性质, 故只涉及定理 5.98(a) 及定理 4.49 中关于参数半变分不等式 (HVI) 的解映射的上导数计算.

推论 5.99(由具有复合势的 HVI 控制 EPEC 的线性次优性)　在定理 5.98(a) 的假设下, 设 $Q(y) = \partial(\psi \circ g)(y)$, 这里

$$S(x) := \big\{ y \in Y \,\big|\, 0 \in q(x, y) + \partial(\psi \circ g)(y) \big\},$$

$q\colon X \times Y \to Y^*$, $g\colon Y \to W$, $\psi\colon W \to \overline{\mathbb{R}}$, 且 W 是 Banach 空间. 此外, 假设 $g \in \mathcal{C}^1$ 有满射导数 $\nabla g(\bar{y})$, 假设 $\nabla g(\cdot)$ 在 $\bar{y}$ 是严格可微的, $\partial\psi$ 的图在 $(\bar{w}, \bar{v})$ 附近是局部闭的, 其中 $\bar{w} := g(\bar{y})$, 而 $\bar{v} \in W^*$ 是满足

$$-q(\bar{x}, \bar{y}) = \nabla g(\bar{y})^* \bar{v}$$

的唯一泛函. 则 $(\bar{x}, \bar{y})$ 相对于 (f, S, Θ) 是线性次优的当且仅当存在 $z^* \in N(\bar{z}; \Theta) \setminus \{0\}$ 和 (唯一定义的)$u \in Y$, 使得

$$0 = \nabla_x f(\bar{x}, \bar{y})^* z^* + \nabla_x q(\bar{x}, \bar{y})^* u,$$

$$0 \in \nabla_y f(\bar{x}, \bar{y})^* z^* + \nabla_y q(\bar{x}, \bar{y})^* u + \nabla g(\bar{y})^* \partial_N^2 \psi(\bar{w}, \bar{v})(\nabla g(\bar{y}) u).$$

证明　应用定理 5.98(a) 并利用定理 4.49 中对映射 S 的 $D_N^* S(\bar{x}, \bar{y})$ 的计算, 该计算基于定理 1.127 中的二阶次微分公式.　△

本小节最后给出由具有复合域的参数广义方程控制的 EPEC 的线性次优性的一个准则.

推论 5.100(由 HVI 控制具有复合域的 EPEC 的线性次优性)　在定理 5.98(a) 的假设下, 设 $Q(y) = (\partial\psi \circ g)(y)$, 这里, 对某 Banach 空间 W 有 $P = W^*$,

$$S(x) := \big\{ y \in Y \,\big|\, 0 \in q(x, y) + (\partial\psi \circ g)(y) \big\},$$

其中 $g\colon Y \to W$, $\psi\colon W \to \overline{\mathbb{R}}$, 且 g 在 $\bar{y}$ 是严格可微的, 有满射导数 $\nabla g(\bar{y})$. 记 $\bar{w} := g(\bar{y})$, $\bar{p} := -q(\bar{x}, \bar{y})$. 假设 $\partial\psi$ 的图在 $(\bar{w}, \bar{p})$ 附近是局部闭的 (当 ψ 是连续的或是顺从的时这个假设自动满足). 则 $(\bar{x}, \bar{y})$ 相对于 (f, S, Θ) 是线性次优的当且仅当存在 $z^* \in N(\bar{z}; \Theta) \setminus \{0\}$ 和 (唯一定义的)$u \in W^{**}$ 满足

$$0 = \nabla_x f(\bar{x}, \bar{y})^* z^* + \nabla_x q(\bar{x}, \bar{y})^* u,$$

$$0 \in \nabla_y f(\bar{x},\bar{y})^* z^* + \nabla_y q(\bar{x},\bar{y})^* u + \nabla g(\bar{y})^* \partial_N^2 \psi(\bar{w},\bar{p})(u).$$

证明 应用定理 5.98(a) 并利用命题 4.53 中对映射 S 的 $D_N^* S(\bar{x},\bar{y})$ 的计算, 该计算基于定理 1.66 中上导数的链式法则. △

5.4.3 极小化问题的线性次优性

本小节研究通常极小化问题的线性次优性; 此时亦称为线性次极小性. 当然, 极小化问题形式是 5.4.2 小节研究的多目标最优化问题在单值 (实值) 目标 f 和 $\Theta = \mathbb{R}_-$ 时的一个特殊子类. 另一方面, 这些问题及其线性次优解与一般多目标问题相比有一些特殊的性质. 下面给出无约束和约束问题近似和点基两种形式的线性次极小性刻画结果, 也将给出一些引人注目的例子. 除了利用下次梯度得到线性次极小性的充要条件之外, 还利用上次梯度得到精细的必要条件, 它们是极小化问题特有的.

定义 5.101(线性次极小性) 设 $\Omega \subset X$, $\varphi\colon X \to \overline{\mathbb{R}}$ 在 $\bar{x} \in \Omega$ 有限. 称 $\bar{x}$ 相对于 (φ, Ω) 是线性次极小的, 如果有

$$\limsup_{x \xrightarrow{\Omega} \bar{x},\, \varphi(x) \to \varphi(\bar{x}),\, r \downarrow 0} \; \inf_{u \in B_r(x) \cap \Omega} \frac{\varphi(u) - \varphi(x)}{r} = 0.$$

在上述定义中如果 $\Omega = X$, 则点 $\bar{x}$ 称为线性次极小的.

值得注意的是, $\bar{x}$ 相对于 (φ, Ω) 的线性次极小性对应于当 $f(x) = \varphi(x) - \varphi(\bar{x})$ 和 $\Theta = \mathbb{R}_-$ 时定义 5.91 中 $\bar{x}$ 相对于 (f, Ω, Θ) 的线性次优性.

易见, 函数 φ 受约束于 $x \in \Omega$ 的任何局部极小点都是相对于 (φ, Ω) 的线性次极小的, 但反之不然. 下面的例子说明了这些概念的一些引人注目的区别, 而这即使在一维无约束问题上也会发生.

例 5.102(线性次极小性的特殊性质) 根据定义可直接验证 $\bar{x} = 0 \in \mathbb{R}$ 对 $\varphi(x) := x^2$, $\varphi(x) := -x^2$ 和 $\varphi(x) := x^3$ 中的每个函数来说都是线性次极小的. 从极小化的角度看, 这些函数的不同之处在于 $\bar{x} = 0$ 分别是它们的极小值点、极大值点和平稳点 (驻点).

点 $\bar{x} = 0$ 也是分段常数和 l.s.c. 函数

$$\varphi(x) := \begin{cases} -\dfrac{1}{n}, & \text{若 } -\dfrac{1}{n} < x \leqslant -\dfrac{1}{n+1}, \quad n \in \mathbb{N}, \\ 0, & \text{若 } x = 0, \\ \dfrac{1}{n}, & \text{若 } \dfrac{1}{n+1} < x \leqslant \dfrac{1}{n}, \quad n \in \mathbb{N} \end{cases}$$

的线性次极小点. 尽管 $\bar{x} = 0$ 对 φ 来说不是局部极小的, 但是 $\bar{x} = 0$ 的任何邻域都

包含这个函数的局部极小值点. 然而, 对 l.s.c. 分段线性函数

$$\varphi(x) := \begin{cases} -\dfrac{1}{n} + \dfrac{1}{n^2}\left(x + \dfrac{1}{n}\right), & \text{若} -\dfrac{1}{n} < x \leqslant -\dfrac{1}{n+1}, \quad n \in \mathbb{N}, \\ 0, & \text{若}\, x = 0, \\ \dfrac{1}{n} + \dfrac{1}{n^2}\left(x - \dfrac{1}{n+1}\right), & \text{若}\, \dfrac{1}{n+1} < x \leqslant \dfrac{1}{n}, \quad n \in \mathbb{N} \end{cases}$$

来说则不然, 它根本就没有局部极小值点, 而 $\bar{x} = 0$ 对它而言却是线性次极小的.

下面在本原空间中给出线性次极小性的一些等价描述, 来说明它与扰动极小性以及利用极限斜率而不是经典导数定义的广义平稳点之间的关系. 考虑到 φ 是增广实值的, 简单起见, 在定义 5.101 中令 $\Omega = X$.

定理 5.103(线性次极小性的等价描述)　设 $\varphi: X \to \overline{\mathbb{R}}$ 在 $\bar{x}$ 有限且在该点附近是 l.s.c. 的, X 是 Banach 空间. 则下面性质等价:

(a) 点 $\bar{x}$ 是 φ 的线性次极小点;

(b) 对任意的 $\varepsilon_k \downarrow 0$, 存在 $x_k \xrightarrow{\varphi} \bar{x} \quad (k \to \infty)$, 使得

$$\varphi(x_k) \leqslant \varphi(x) + \varepsilon_k \|x - x_k\|$$

对任意 x_k 附近的 x 和 $k \in \mathbb{N}$ 成立.

(c) 有

$$\liminf_{x \xrightarrow{\varphi} \bar{x}} |\ \nabla\varphi\ |(x) = 0,$$

这里

$$|\nabla\varphi|(x) := \limsup_{u \xrightarrow{\varphi} x} \frac{\max\{\varphi(x) - \varphi(u), 0\}}{\|u - x\|}$$

称为 φ 在 x 处的 (强) 斜率.

证明　易见 (b) 中的性质可等价地描述为

$$\limsup_{x \xrightarrow{\varphi} \bar{x},\, r \downarrow 0} \tau_\varphi(x, r) = 0, \quad \text{其中} \quad \tau_\varphi(x, r) := \inf_{\|u - x\| < r} \min\left\{\frac{\varphi(u) - \varphi(x)}{\|u - x\|}, 0\right\}.$$

由于

$$\tau_\varphi(x, r) \leqslant \inf_{\|u - x\| \leqslant r} \frac{\varphi(u) - \varphi(x)}{r} \leqslant 0, \quad \forall r > 0.$$

则 (b)⇒(a). 为证 (a) ⇒(b), 假设 $\bar{x}$ 是 φ 的线性次极小点. 根据定义 5.101,

找到序列 $u_k \xrightarrow{\varphi} \bar{x}$ 和 $\varepsilon_k \downarrow 0$ 满足

$$\varphi(x) - \varphi(u_k) \geqslant -\varepsilon_k^2, \quad \forall x \in u_k + 2\varepsilon_k \mathbb{B}.$$

根据定理 2.26 中的 Ekeland 变分原理, 对任意 $k \in \mathbb{N}$, 都存在 $x_k \in u_k + \varepsilon_k \mathbb{B}$ 满足

$$\varphi(x_k) \leqslant \varphi(u_k), \quad \varphi(x) - \varphi(x_k) \geqslant -\varepsilon_k \|x - x_k\|$$

对 x_k 附近的任意 x 成立. 显然有 $x_k \xrightarrow{\Omega} \bar{x}$ 和 $\tau_\varphi(x_k, r) \geqslant -\varepsilon_k$ 对小的 $r > 0$ 成立, 根据 (b) 的上述描述, 这蕴涵着性质 (b).

最后注意到

$$\limsup_{x \xrightarrow{\mathcal{L}} \bar{x},\, r \downarrow 0} \tau_\varphi(x, r) = -\liminf_{x \xrightarrow{\mathcal{L}} \bar{x}} |\nabla\varphi|(x),$$

可得到等价关系 (b)⇔ (c), 从而完成了定理的证明. △

标准局部极小值点与线性次极小点的最主要的区别之一在于前者相对于小的扰动不是稳定的, 而后者是稳定的. 实际上, 考虑最简单的二次函数 $\varphi(x) := x^2$, $\bar{x} = 0$ 是它的全局极小点. 在 $\bar{x}$ 附近用 $\psi(x) := -|x|^{3/2}$ 扰动 φ, 可看到 $\bar{x}$ 不再是函数 $\varphi(x) + \psi(x)$ 的局部极小点; 事实上, 这个函数在 $\bar{x} = 0$ 达到它的全局最大值. 另一方面, 线性次极小性相对于导数为零的任何光滑扰动来说都是稳定的.

命题 5.104(线性次极小性的平稳性) 设 $\Omega \subset X$, 设 $\varphi\colon X \to \overline{\mathbb{R}}$ 和 $\psi\colon X \to \mathbb{R}$ 是 Banach 空间 X 上的函数, 满足 ψ 在 $\bar{x}$ 是严格可微的且 $\nabla\psi(\bar{x}) = 0$. 则 $\bar{x}$ 相对于 (φ, Ω) 是线性次极小的当且仅当它相对于 $(\varphi + \psi, \Omega)$ 是线性次极小的.

证明 由定义 5.101 及对 ψ 的严格导数 $\nabla\psi(\bar{x}) = 0$ 直接可以得到 $\bar{x}$ 相对于 (φ, Ω) 的线性次极小性蕴涵 $\bar{x}$ 相对于 $(\varphi + \psi, \Omega)$ 的线性次极小性. 对函数 $\varphi + \psi$ 和 $-\psi$ 应用这个结果, 则有相反的蕴涵关系成立. △

这个结果的一个直接推论是在任意 Banach 空间中, 任意光滑函数的线性次极小点简化为传统意义下它的平稳点.

推论 5.105(严格可微函数的线性次极小点和平稳点) 设 $\varphi\colon X \to \mathbb{R}$ 在 $\bar{x}$ 是严格可微的. 则 $\bar{x}$ 是 φ 的线性次极小点当且仅当 $\bar{x}$ 是 φ- 平稳点, 即有 $\nabla\varphi(\bar{x}) = 0$.

证明 当 $\varphi = 0$, $\Omega = X$ 时, 由命题 5.104 可得. △

由推论 5.105 可知, 对 Banach 空间上的严格可微实值函数来说, 线性次极小性与对称的线性次极大性的概念是等价的. 然而, 对非光滑函数而言则不然. 因此, 线性次极小性与线性次极大性这两个概念可看做是非光滑函数传统平稳点 (驻点) 概念的单边推广.

现基于定理 5.92 和定理 5.94 建立线性次极小性的模糊/近似和点基/确切形式的充要条件. 简洁起见, 只给出这个性质的准则而不是分开的必要条件和充分条件. 下面的定理含有这个方向在一般情况下利用费用函数 φ 在约束集 Ω 上的限制 $\varphi_\Omega(x) := \varphi(x) + \delta(x; \Omega)$ 的 Fréchet 和基本次梯度表示的压缩结果.

定理 5.106(线性次极小性压缩型次微分准则) 设 $\varphi_\Omega\colon X \to \overline{\mathbb{R}}$ 在 $\bar{x} \in \Omega$ 附近是 l.s.c. 的且 $|\varphi(\bar{x})| < \infty$, 设 X 是 Asplund 空间. 则下列结论成立:

(i) 点 $\bar{x}$ 相对于 (φ, Ω) 是线性次极小的当且仅当对任意的 $\varepsilon > 0$, 存在 $x \in \Omega \cap (\bar{x} + \varepsilon\mathbb{B})$ 满足 $|\varphi(x) - \varphi(\bar{x})| \leqslant \varepsilon$ 且存在 $x^* \in \partial\varphi_\Omega(x)$ 满足 $\|x^*\| \leqslant \varepsilon$.

(ii) 假设 $\dim X < \infty$. 则 $\bar{x}$ 相对于 (φ, Ω) 是线性次极小的当且仅当 $0 \in \partial\varphi_\Omega(\bar{x})$.

证明　(i) 由当 $\Theta = \mathbb{R}_-$ 和 $f(x) = \varphi(x) - \varphi(\bar{x})$ 时定理 5.92 中线性次优性的模糊刻画可得.

为证 (ii), 对与 (i) 中相同的 f 和 (5.111) 中定义的 F 应用定理 5.94(iii) 中线性次优性的点基刻画. 注意到, 由于 $Z = \mathbb{R}$, 这个 F 在 $(\bar{x}, 0)$ 处自动是 SNC 且强上导数正规的, 并且显然有

$$0 \in D^*F(\bar{x}, 0)(1) \Longleftrightarrow 0 \in \partial\varphi_\Omega(\bar{x}).$$

这就完成了定理的证明. △

值得注意的是定理 5.106(i) 中的 ε- 次微分条件不能用 $0 \in \partial\varphi_\Omega(x)$ 替代; 一个反例可由例 5.102 中的第二个函数给出.

定理 5.106 的第二个结论和等式型次微分和法则蕴涵着下面的结果, 这个结果给出了用在参考解 $\bar{x}$ 处所计算的 φ 的基本次梯度和 Ω 的基本法锥表示的线性次极小性的点基刻画.

推论 5.107(线性次极小性分离型的点基刻画)　设 $\dim X < \infty$, $\bar{x} \in \Omega$ 且 $|\varphi(\bar{x})| < \infty$. 还假设下列假设 (a)~(c) 之一成立:

(a) φ 在 $\bar{x}$ 附近是 l.s.c. 的且 $\Omega = X$.

(b) φ 在 $\bar{x}$ 是严格可微的且 Ω 在该点附近是闭的.

(c) φ 在 $\bar{x}$ 附近是 l.s.c. 的且在该点是下正则的, Ω 在 $\bar{x}$ 是局部闭和法向正则的, 且有规范条件

$$\partial^\infty\varphi(\bar{x}) \cap \big(-N(\bar{x}; \Omega)\big) = \{0\}.$$

则 $\bar{x}$ 相对于 (φ, Ω) 是线性次极小的当且仅当

$$0 \in \partial\varphi(\bar{x}) + N(\bar{x}; \Omega). \tag{5.114}$$

证明　当 $\Omega = X$ 时条件 (5.114) 与定理 5.106(ii) 中的条件相同. 当 φ 是严格可微的, 条件 $0 \in \partial\varphi_\Omega(x)$ 根据命题 1.107(ii) 中的等式

$$\partial\varphi_\Omega(\bar{x}) = \nabla\varphi(\bar{x}) + N(\bar{x}; \Omega)$$

等价于 (5.114). 在 (c) 的假设下有等式

$$\partial\varphi_\Omega(\bar{x}) = \partial\varphi(\bar{x}) + N(\bar{x}; \Omega),$$

这源于定理 3.36 中的等式和法则. △

值得注意的是, 情形 (b) 下线性次极小性的刻画 (5.114) 由关于多目标最优化的定理 5.95(a) 直接可得, 而当 φ 是局部 Lipschitz 时, 情形 (a) 中的结果由定理 5.95(b) 可得. 然而, 情形 (c) 下根据推论 5.107 保证 (5.114) 成立的假设本质上弱于根据定理 5.95(c) 而来的假设. 事实上, 定理 5.95(c) 中当 $Z = \mathbb{R}$ 时关于 $f(x) = \varphi(x) - \varphi(\bar{x})$ 的 N- 正则性假设实际上是 φ 在 $\bar{x}$ 的图正则性, 根据命题 1.94 等价于 φ 在该点的严格可微性. 另一方面, 推论 5.107(c) 中假设的 φ 的下正则性对极小化问题中出现的非常重要的非光滑函数类成立. 特别地, 包含凸函数和上面讨论的更广泛的一类顺从函数. 在极小化问题的情形下, 定理 5.95 的结果与推论 5.107 的结果之间这样的差别归因于最小化增广实值函数的单边特性, 在向量框架下这样的差别在分离条件中不复存在.

在 $\Omega \neq X$ 的约束情形下, 根据推论 5.107 的结果, 可得到特殊约束型问题线性次极小性充要条件的结果. 对算子、泛函和/或均衡约束问题 (即 MPEC) 可得到类似推论 5.96, 推论 5.97, 定理 5.98 及其两个推论的结果. 而且, 除了要求目标映射的严格可微性的上述结果之外, 还可得到具有正则约束和下正则费用函数的问题的线性次极小性的刻画. 具体细节留给读者.

本小节的最后得到非光滑约束问题线性次极小性的必要条件, 其中上次梯度被用于描述单目标和不等式约束的函数.

考虑在 $\bar{x}$ 有限的费用函数 $\varphi_0\colon X \to \overline{\mathbb{R}}$ 和由

$$\Delta := \{x \in \Omega \subset X \ \text{ 且 } \ \varphi_i(x) \leqslant 0,\ i = 1, \cdots, m\}$$

给出的约束集合 $\Delta \subset X$, 这里 $\varphi_i\colon X \to \mathbb{R}$ $(\forall i)$. 下一定理给出了相对于 (φ_0, Δ) 的线性次极小解的上次微分必要条件.

定理 5.108(线性次极小解的上次微分必要条件) 设 $\bar{x} \in \Delta$ 相对于 (φ_0, Δ) 是线性次极小的, 其中 Ω 在 $\bar{x}$ 附近是局部闭的. 假设或者 X 有一 Lipschitz $\mathcal{C}^1$ 阻尼函数, 或者 X 是 Asplund 空间且 $\varphi_i(\bar{x}) < 0$ $(i = 1, \cdots, m)$. 则对任意的 Fréchet 上次梯度 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$ $(i = 0, \cdots, m)$, 存在 $0 \neq (\lambda_0, \cdots, \lambda_m) \in \mathbb{R}^{m+1}$ 使得

$$\lambda_i \geqslant 0 \ (i = 0, \cdots, m), \quad \lambda_i\varphi_i(\bar{x}) = 0 \quad (i = 1, \cdots, m),$$

$$-\sum_{i=0}^{m} \lambda_i x_i^* \in N(\bar{x}; \Omega).$$

证明 假设 $\widehat{\partial}^+\varphi_i(\bar{x}) \neq \varnothing$ $(i = 0, \cdots, m)$ (否则定理的结论平凡成立), 对每个 i 任取 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x})$. 现应用定理 1.88(i) 中的 Fréchet 次梯度的变分描述于 $-x_i^* \in \widehat{\partial}(-\varphi_i)(\bar{x})$, 找到函数 $s_i\colon X \to \overline{\mathbb{R}}$ $(i = 0, \cdots, m)$, 在 $\bar{x}$ 处 Fréchet 可微且满足

$$s_i(\bar{x}) = \varphi_i(\bar{x}), \quad \nabla s_i(\bar{x}) = x_i^* \text{和} \ \ s_i(x) \geqslant \varphi_i(x) \ \text{ 在 } \bar{x} \text{ 附近成立.}$$

考虑另一个约束集

$$\widetilde{\Delta} := \{x \in \Omega \ \text{ 且 } \ s_i(x) \leqslant 0 \ \ (i = 1, \cdots, m)\}.$$

并注意到 $\bar{x} \in \widetilde{\Delta}$ 且 $\bar{x}$ 相对于 $(\varphi_0, \widetilde{\Delta})$ 是线性次极小的. 而且, 根据 s_0 的构造以及线性次极小性和 Fréchet 上次梯度的定义知 $\bar{x}$ 相对于 $(s_0, \widetilde{\Delta})$ 是线性次极小的. 如果 $s_i(\bar{x}) = \varphi_i(\bar{x}) < 0$ 对所有 $i = 1, \cdots, m$ 成立, 当 $f(x) = \varphi(x) - \varphi(\bar{x}), \Theta = \mathbb{R}_-$ 时, 根据推论 5.97(其中的必要性部分在任意 Asplund 空间显然成立, 参看定理 5.94(ii) 和随后基于 Asplund 空间中分析法则的论证), 有

$$-\nabla s_0(\bar{x}) = -x_0^* \in N(\bar{x}; \widetilde{\Delta}) = N(\bar{x}; \Omega).$$

余下考虑当至少一个不等式约束在 $\bar{x}$ 起作用时定理的另一种情形. 此时当 $\mathcal{S} = \mathcal{LC}^1$ 时, 根据定理 1.88(ii), 所有函数 s_i 可选为在 $\bar{x}$ 附近是连续可微的. 于是再次应用在 Asplund 空间成立的推论 5.97 中线性次极小性的必要条件, 得包含关系

$$-\sum_{i=0}^{m} \lambda_i \nabla s_i(\bar{x}) \in N(\bar{x}; \Omega)$$

对某 $(\lambda_0, \cdots, \lambda_m) \neq 0$ 成立且满足上述符号和互补松弛条件. 定理中的最后一个关系现由 $\nabla s_i(\bar{x}) = x_i^* \ (i = 1, \cdots, m)$ 可得. △

当具体确定约束集 Ω 是等式、算子、均衡和/或其他类型的约束形式时, 应用十分完善的分析法则, 与在 5.1 节和 5.2 节中建立的极小化问题的上次微分必要条件类似, 根据定理 5.108 可得涉及费用函数 Fréchet 上次梯度的线性次优性的必要条件.

5.5 第 5 章的评注

5.5.1 分析和最优化之间的双边关系

本章主要研究第 I 卷 (1~4 章) 所建立的变分分析的基本工具在最优化和均衡问题中的应用. 更具体地说, 本章主要研究各种非动态约束最优化问题 (包括均衡约束优化问题) 和多目标优化问题, 涵盖了均衡的经典和广义的概念. 主要致力于利用在上面建立起来的基本极点/变分原理和广义微分 (及其广泛的分析法则) 的工具, 得到所研究问题的各种类型的必要最优性和次优性条件.

众所周知, 最优化/变分思想和技术在数学分析的所有领域中起着至关重要的作用, 包括那些似乎与最优化很遥远的领域. 回顾在前言中提到的引人注目的例子, 最起初的 (Fermat) 导数概念, 它的引入是受一个最优化问题的促动; 经典的 (Lagrange) 中值定理, 可能是微积分学中最基本的结果, 它的证明就是基于化归为

最优化和 Fermat 驻点原理; 还有 Bernoulli 的最速降线问题, 实际上催生了整个 (无限维) 泛函分析的发展.

最优化的巨大作用的另一个有力的例证是在上卷中建立起来的广义微分分析法则, 它是以变分思想为依据, 主要基于极点原理. 我们知道, 极点原理给出了集合极性的必要条件, 而集合极性可看作最优性的几何概念, 推广了各种与最优化相关和均衡问题最优解的经典和广义的概念. 因此, 在约束最优化和均衡的具体情形中, 极点原理的应用和具体化直接给出了这种情形下的必要最优性条件. 然而, 当利用无限维中广义微分分析法则以及相关联的 SNC 分析法则的力量时, 则可得到更多先进和不同的结果. 这就是第 5 章的主要内容.

值得一提的是, 基于极点原理获得的必要最优性条件的方法, 以及极点原理本身及其证明, 本质上区别于针对约束最优化必要最优性条件的传统方法, 该方法由 Dubovitskii 和 Milyutin[369, 370] 提出和形式化, 接着在随后许多的文献中得到发展; 参见 1.4.1 小节的一些引文和讨论. Dubovitskii-Milyutin 形式体系包括下面三个主要组成部分:

(a) 通过本原/初始空间中架构于初始费用和约束数据上的特定集合的交集为空的性质来处理局部极小值;

(b) 利用无交集的凸锥来逼近上面的集合;

(c) 利用凸分离来得到抽象 Euler 方程形式的对偶必要最优性条件.

本书的极点原理方法与 Dubovitskii 和 Milyutin 形式体系的基本区别在于在本原空间中没有进行任何凸逼近, 而广义 Euler 方程是通过约化为光滑无约束优化问题组成的逼近序列在对偶空间中利用非凸结构直接得出的; 参见第 2 章.

5.5.2 非光滑分析和最优化中的下和上次梯度

对增广实值函数的极小化问题, 本书在所给的结果中区别下次梯度和上次梯度最优性条件. 这两种条件对非光滑费用函数情形而言是明显不同的, 当然对如命题 5.1 中的光滑目标函数来说, 这两种条件相同. 命题 5.1 类型的第一个结果是针对可以由凸锥 K 逼近的集合 $\Omega \subset X$ 的, 由 Kantorovich[664] 早在 1940 年于一般拓扑空间 X 中利用 K 的对偶/共轭锥 K^* 以

$$-\nabla\varphi(\bar{x}) \in K^*$$

形式得到. Kantorovich 的文章是俄文的, 其结果可能是极点问题一般理论的第一个结果. 遗憾的是, 这个结果实在是太超前了, 无论在前苏联还是在西方都没有引起任何关注. 建议读者参阅 Polyak 的文章 [1099] 对此结果和关于最优化其他早期发展成果的出色分析, 包括前苏联当时相关的社会环境.

在非光滑最优化中, 次梯度 (或作为次梯度集合的次微分) 的概念传统上与非

光滑函数的“下方”性质有关, 从而与极小化 (而非极大化) 问题有关. 另一方面, 凹函数的次梯度/次微分由 Rockafellar[1142] 用不同于 (然而对称与) 凸函数的次梯度/次微分的方法定义. 实际上它对应于现在所称的上次梯度/次微分; 这一术语由 Rockafellar 和 Wets[1165] 正式引入, 尽管上次梯度结构实际上在该书中并没有用到.

另一个在非线性偏微分方程黏性解理论和许多关于非光滑分析的著作中已被完全接受的术语是“次微分”和“超微分”(或上微分, superdifferential). 有意思的是, (下) 次梯度被用于定义黏性“超解”(supersulution), 而“下解”(subsolution) 却通过“超梯度”定义. 经过和 Rockafellar 与 Wets 讨论之后, 本书选择使用“下”和“上”次梯度 (lower subgradient, upper subgradient) 术语, 这对最优化来说更自然、更适合, “下”(lower) 默认地来描述次微分结构 (“下”一般省略), 推广了凸函数的相应结构, 并且对对称结构应用“上”(upper) 替代“超”(super), 推广了凸分析框架下凹函数的相应结构.

在此方向上值得回顾的是关于局部 Lipschitz 函数类的 Clarke 广义梯度 (这个术语很好), 作为下次微分结构 (推广了凸而不是凹函数的次微分), 由于它加–减号的对称性: $\partial_C(-\varphi)(\bar{x}) = -\partial_C\varphi(\bar{x})$, 同时与它的上次微分结构相同, 这与经典梯度相类似. 特别地, 这蕴涵着通过 Clarke 广义梯度描述的任何条件不区分非光滑 (甚至凸) 函数的极小化和极大化, 不区分相反不等式符号的不等式约束等. 然而, 正如 Rockafellar[1142] 描述的那样, “凸函数相对于凸集极大化的理论有完全不同于极小化理论的特征”. 与 Clarke 的结构不同, 本书的下和上类 Fréchet 和基本/极限次微分结构本质上是单边的, 并且各不相同. 通过有效地利用这些差别, 第 5 章中给出了非光滑函数约束极小化的下和上次微分最优性条件.

5.5.3　凸函数及凸函数的差的极大化问题

就我们所知, 区分极大化和极小化的第一个必要最优条件由Rockafellar[1142, 32 节] 在有限维空间的凸集 Ω 上极大化凸函数 φ 的问题而得. 对局部极大值点 $\bar{x} \in \Omega$, 这个条件由集合包含的形式

$$\partial\varphi(\bar{x}) \subset N(\bar{x};\Omega) \tag{5.115}$$

给出. 对在 Ω 上极小化凹函数 $-\varphi$ 的问题来说, 这个条件显然简化为命题 5.2 中的两个包含关系 (5.3). 正如 5.1.1 小节所述, 有非常重要的一类 DC 函数, 即两个凸函数的差 $\varphi_1 - \varphi_2$, 可被简化为在凸集上极小化凹函数. 对 DC 函数来说, 与必要条件 (5.115) 类似的条件记为

$$\partial\varphi_1(\bar{x}) \subset \partial\varphi_2(\bar{x}); \tag{5.116}$$

参见文献 [573]. 那个时候 (5.115) 和 (5.116) 的一些修改版本被用于得到凸函

数、DC 函数和与它们密切相关的函数在凸集上的全局极大化的充要条件; 特别地, 参见 Strekalovsky [1226–1228], Hiriart-Urruty [573], Hiriart-Urruty 与 Ledyaev [574], Flores-Bazán[461], Flores-Bazán 与 Oettli[462], 和 Tsevendorj [1272]. 读者在 Dür, Horst 与 Locatelli 的综述文章 [373] 和 Ernst 与 Théra 的近期研究 [410] 中可找到关于这个方向主要成果的更多细节和讨论, 其中极大化和极小化凸函数之间其他一些引人注目的差别也被发现. 还建议读者参看 Dutta 的近来研究 [375], 他利用 Clarke 广义梯度在有限维空间的凸集上得到若干类“伪凸”和“拟凸”函数全局极大值点的刻画. 而且, 他在这样的集合上利用我们的基本次微分结构得到了一般 Lipschitz 函数全局极大化的充分条件.

5.5.4 约束极小化的上次微分条件

关于一般（可以是非 Lipschitz 的）费用函数的约束极小化问题上次微分条件的系统研究见 Mordukhovich[925] 在无限维空间中的工作. 第 5 章给出的这种类型的大多数结果取自该文. 所得结果即使在有限维空间中似乎也是新的. 这些条件应用于局部极小值点, 与更传统的下次微分条件类似, 在第 5 章中用平行的方式给出. 如前述, 这两种必要最优性条件一般来说是独立的, 然而上次微分必要最优性条件对涉及非光滑费用函数 φ 的若干类极小化问题来说有时本质上可以更强些. 虽然关系 $\widehat{\partial}^+\varphi(\bar{x}) = \varnothing$ 本身对局部极小点 $\bar{x}$ 来说是易于验证的必要条件, 但上次微分最优性条件最有效的应用需要 Fréchet 上次微分的非平凡性 $\widehat{\partial}^+\varphi(\bar{x}) \neq \varnothing$. 特别地, 对所讨论问题中在极小值点恰好是上正则的 Asplund 空间上的那些局部 Lipschitz 函数而言, 这个条件自动成立; 更多细节参见注 5.4. 后面这类函数除了光滑和凹连续函数之外还包括一大类半凹函数, 它们在各种应用中非常重要, 尤其在最优控制和非线性 PDE 的黏性解理论中.

回顾定义在凸集 Ω 上的函数 $\varphi\colon \Omega \to \mathbb{R}$ 是半凹的, 如果存在非递减上半连续函数 $\omega\colon \mathbb{R}_+ \to \mathbb{R}_+$ 满足 $\omega(\rho) \to 0\ (\rho \downarrow 0)$ 使得

$$
\begin{aligned}
&\lambda\varphi(x_1) + (1-\lambda)\varphi(x_2) - \varphi\big(\lambda x_1 + (1-\lambda)x_2\big) \\
&\leqslant \lambda(1-\lambda)\|x_1 - x_2\|\omega\big(\|x_1 - x_2\|\big)
\end{aligned}
\tag{5.117}
$$

对任意 $x_1, x_2 \in \Omega, \lambda \in [0,1]$ 成立; 参见 Cannarsa 与 Sinestrari 近来的书 [217] 及其中的参考文献. 半凹性的理论和应用中最重要的情形对应于 (5.117) 中的线性模 $\omega(\cdot)$. 这类半凹函数（用等价的形式和相反的符号）可能首次由 Janin[629] 在“平方凸性”(或法文“Presque Convexes du deuxieme ordre”, PC2) 的名称下在最优化中引入和使用. 然而, 这个结构可追溯到偏微分方程理论, 这个半凹 (具有线性模) 函数类正是由 Kruzhkov[720] 和 Douglis[368] 使用的函数类, 用来建立关于 Hamilton-Jacobi 方程的解的第一个全局存在和唯一性结果. 而且, 半凹函数在

Hamilton-Jacobi 和类似方程广义 (黏性、极小极大等) 解的强有力的唯一性理论及其在最优控制和微分对策的应用中起着不同寻常的作用; 特别地可参见文献 [85, 86, 216, 217, 295–297, 458, 471, 472, 789, 793, 1230] 及其参考文献.

有意思的是, 半凹函数与次光滑, 或上 -$\mathcal{C}^k$ 函数类的主要子类之间有密切的关系 (实际上是等价性关系). 次光滑函数由 Rockafellar[1151] 引入, 即形式为

$$\varphi(x) := \min_{t \in T} \phi(x,t)$$

的函数, 其中 T 是紧空间, $\phi(x,t)$ 在一开集上的任意点 $x \in \mathbb{R}^n$ 处关于 $t \in T$ 是一致地 k 次连续可微的; 关于某些无限维的推广也可参见 Penot 的文章 [1069]. 正如 Rockafellar[1151, 1165] 所证, 上 -$\mathcal{C}^2$ 类函数完全与具有线性模的半凹函数类, 即平方凹的函数类相同. 一般半凹函数类 (5.117) 与上 -$\mathcal{C}^1$ 函数类之间的等价性由 Cannarsa 与 Sinestrari[217] 建立. 而且, 上 -$\mathcal{C}^2$ 函数恰好等价于 Vial[1286] 意义下的“弱凹”函数, 而上 -$\mathcal{C}^1$ 函数 (在有限维中) 与 Ngai, Luc 与 Théra[1006] 所研究的“近似凹”函数相同. 关于这些及相关非光滑函数类的更多讨论和相关联的几何概念的广泛研究, 建议读者参见 Aussel, Daniilidis 与 Thibault 的近期文章 [63].

还注意到具有线性和更一般模的半凹函数密切关联于广义凸性理论中所谓的仿凹函数; 参见 [534, 697, 1040, 1072] 及其参考文献. 这个名称由 Rolewicz[1169, 1170] 提出, 他在集值映射的框架下独立地引入和研究了仿凸性/仿凹性. 对这样函数的强烈兴趣还源于通过下卷积和相关算子进行的逼近和正则化过程, 在许多重要情形中已被证明是局部 $\mathcal{C}^{1,1}$ 的, 这源于由 Hiriart-Urruty 与 Plazanet[576] 首次建立的如下刻画：一函数在 $\bar{x}$ 附近是 $\mathcal{C}^{1,1}$ 的, 如果它在该点附近同时是仿凸和仿凹的. 特别地, 关于这个结果在二阶广义微分中的各种应用, 建议读者参看 Eberhard 等的文章 [381, 386, 387].

5.5.5 约束极小化的下次微分最优性和规范条件

与非光滑极小化的上次微分条件相比, 下次微分条件更常用, 具有各种各样的变种, 并且有更长的历史. 当然, 对具有光滑数据的最优化问题来说, 下和上次微分必要最优性条件简化为约束最优化的经典结果, 可追溯到规范 (有的称为正常的或 Karash-Kuhn-Tucker) 和非规范 (有时称为 Fritz John) Lagrange 乘子的标准版本. 第一种类型的结果包含规范条件, 来保证相应于目标/费用函数乘子的非平凡性 ($\lambda_0 \neq 0$). 关于这样最优性和规范条件的起源和动因, 建议读者参看奠基性文献 Lagrange[737], Karush[665](在 Kuhn[723] 的综述文章中发表), John[638], Kuhn 与 Tucker[724] 和 Mangasarian 与 Fromovitz[841]. 进一步的发展、更详细的历史描述和各种应用特别地可在下列文献及其参考文献中找到：[7, 9, 89, 111, 112, 158, 163, 164, 249, 255, 370, 376, 432, 499, 504, 512, 544, 571, 588, 595, 602, 618, 707, 718, 801,

824, 860, 840, 892, 902, 962, 1009, 1097, 1119, 1152, 1155, 1160, 1165, 1216, 1256, 1264, 1265, 1267, 1268, 1289, 1315, 1319, 1340, 1341, 1373, 1378].

值得注意的是, 第 5 章给出的最优化的条件与第 I 卷从广义微分分析法则的角度所得的规范条件有相同的本质, 是密切相关的. 而且, 本书的最优性与规范条件都在对偶空间中得到, 一般来说比它们的本原空间对应物有更少的限制. 这些最优性和规范条件共同的对偶空间结构允许在第 5 章建立的规范和非规范型最优化结果之间建立一座自然的桥梁.

本书主要关注各类优化问题的一阶必要最优性 (和次优性) 条件. 然而, 对具有均衡约束问题, 不但应用一阶而且应用二阶次微分结构, 这是由于这样的约束的一阶变分性质; 参看 5.2 节和 5.3 节和下面给出的关于它们的相应评注. 读者可在下列文献及其参考文献中找到关于二阶最优性条件的更多信息：[37, 64, 65, 102, 111, 132, 133, 153, 176, 234, 236, 282, 283, 372, 384, 387, 502, 575, 486, 516, 601, 613, 624, 628, 704, 756, 764, 771, 857, 858, 877, 1037-1039, 1067, 1092, 1156, 1165, 1307, 1308, 1310, 1337, 1358].

5.5.6 具有算子约束的最优化问题

5.1.2 小节的素材主要关于具有通过集合在极值映射下的逆像/原像以一般形式 $x \in F^{-1}(\Theta) \cap \Omega$ 定义的所谓算子约束的最小化问题下和上次微分两种形式的必要最优性条件. 传统上, 算子约束由等式形式 $f(x) = 0$ 定义, 其中 $f\colon X \to Y$ 是有无限维值空间 Y 的单值映射. 这个名称 (可能首次出现在俄文文献中) 的出现是因为注意到变分法和最优控制的典型问题中的动态约束可写为这样的一种形式, 其中 f 是特定的映入无限维空间内的微分或积分算子; 例如, 参见 Dubovitskii 与 Milyutin 的文章 [370].

似乎对于这样问题 Lagrange 乘子无限维形式的第一个一般结果由 Lyusternik 在他开创性的工作 [824] 中得到, 其中 f 是 Banach 空间上的 $\mathcal{C}^1$ 算子. 为建立这个结果, Lyusternik 发展了现在已成为经典的迭代过程, 并且得到“距离估计”, 现被称为度量正则性. Lagrange 原理的 Lyusternik 版本 (规范形式) 在 Lyusternik 正则性条件 $\ker \nabla f(\bar{x})^* = 0$ 下得到, 意味着导数算子 $\nabla f(\bar{x})\colon X \to Y$ 的满射性. 如果导数像 $\nabla f(\bar{x})X$ 在 Y 中是闭的, 由 Lyusternik 的规范必要最优性条件不难得到关于受约束于算子约束 $f(x) = 0$ 的光滑函数 φ 的局部极小值点 $\bar{x}$ 的 Lagrange 乘子原理的非规范版本:

$$\lambda \nabla \varphi(\bar{x}) + \nabla f(\bar{x})^* y^* = 0, \quad \lambda \geqslant 0, \quad (\lambda, y^*) \neq 0, \tag{5.118}$$

例如, 参见 Ioffe 与 Tikhomirov 的著作 [618]. 众所周知, 如果没有导数像集的闭性假设, 对光滑问题. 即使在 $Y = \ell^2$ 这样最简单的无限维空间情形, 乘子规则 (5.118)

一般来说不成立.

对极小化费用函数 $\varphi_0(x)$, 约束包括非光滑约束 $f(x)=0$, 其中 $f\colon X\to Y$ 为 Banach 空间之间的 Lipschitz 映射, 还包括更标准的约束

$$\varphi_i(x)\leqslant 0,\quad i=1,\cdots,m,\quad x\in\Omega,$$

第一个必要最优性条件由 Ioffe[595] 得到, 其中利用了 Clarke 广义梯度和法锥并且有广义 Lagrange 形式如下:

$$0\in\partial_C\left(\sum_{i=0}^{m}\lambda_i\varphi_i+\langle y^*,f\rangle\right)(\bar{x})+N_C(\bar{x};\Omega),\quad (\lambda_0,\cdots,\lambda_m,y^*)\neq 0, \tag{5.119}$$

当然还有符号和互补松弛条件. 除 φ_i $(i=0,\cdots,m)$ 的传统的局部 Lipschitz 性质之外, 在文献 [595] 中还假设: Y 有等价的范数使它的对偶是严格凸的; Ω 在 $\bar{x}$ 有某种利用 Clarke 切锥和方向导数描述的“切向下半连续性质”; 并且 f 有一“严格预导数 (prederivative)”具有范数紧值满足相对于 $T_C(\bar{x};\Omega)$ 的一种“有限余维性质”. 这个结果由 Ioffe[598] 和 Ginsburg 与 Ioffe[506] 得到了改进, 通过利用“近似”次微分和法锥而不是 Clarke 的凸值结构, 在利用“近似”法锥和次梯度结构描述的有限余维性质的更加精细版本的条件下, 他们建立了明显更强的 (5.119) 的对应结果. 值得注意的是, 上面后一种有限余维性质的先进描述恰好密切关联于映射部分序列法紧性 (PSNC) 的拓扑对应物, 还关联于由 Jourani 与 Thibault[655] 得到的部分 CEL 性质; 参见 1.2.5 小节, 4.4.5 小节和文献 [607] 中相应的讨论.

5.5.7　由基本分析法则处理算子约束

给出算子约束的一般问题上和下次微分两种形式的非规范必要条件的定理 5.11 在文献 [925] 中得到, 而定理 5.7 和定理 5.8 中的规范结果是新的. 注意到, 规范最优性条件实际上蕴涵着相应的非规范最优性条件, 但反之不然. 这是由于定理 5.7 和定理 5.8 (以及本书随后给出的必要最优性条件) 中规范条件的结构, 它们通常包含对偶空间更精细的的信息, 这对非规范条件是不需要的. 还注意到完善的 SNC 分析法则允许得到各种类法紧性条件, 一般来说, 比前面提到的有限余维性质有更少的限制, 保证算子约束的问题点基必要最优性条件成立.

值得一提的是, 本书建立必要最优性条件的方法中把算子约束看成几何约束, 然后通过初始数据利用广义微分和 SNC 分析法则得到结果. 在极点原理的基础上, 利用这两种分析法则 (这是本书所应用的基本结构特有的), 无疑恰好是成功实施我们方法的最关键因素. 值得注意的是, 这个方法对处理更多个几何约束没有任何限制, 这对各类最优化问题有重要意义, 尤其对最优控制; 参见第 6 章和第 7 章. 众所周知, 只允许有一个内部可能是空集的几何约束 (或算子/等式型约束) 是 Dubovitskii-Milyutin 形式体系及其随后发展中的一个重大障碍.

在关于定理 5.7, 定理 5.8 和定理 5.11 的讨论的最后, 评论一下定理 5.7 和定理 5.11(i) 的那些部分, 它们对在有限维空间中取值的等式型约束没有严格 (或连续) 可微性假设. 这些结果本质上源于推论 1.15 给出的逆像的 Fréchet 法锥的计算, 这基于 Brouwer 不动点理论; 请比照 Halkin[543] 和 Ioffe[595]. 这种类型的结果由 Ioffe[595, 602] 和 Ye[1340, 1341] 建立来得到具有有限多等式和不等式约束的 Lipschitz 问题的必要最优性条件, 其中利用没有任何鲁棒性更小的凸值次微分 (Michel-Penot[870, 871] 和 Treiman[1264, 1265] 类型). 注意到, 定理 5.7(i) 和 5.11(i) 中的相应结果不要求约束函数的局部 Lipschitz 连续性, 而在所论点附近仍对等式约束函数加连续性条件. 这个条件对 Lagrange 型必要最优性条件的成立是至关重要的, 正如例 5.12 表明的那样, 该例取材于 Uderzo[1274].

5.5.8 精确惩罚与弱化的度量正则性

5.1.2 小节的剩余部分涉及另一方法来处理关于由 Lipschitz 映射给出的经典等式型 $f(x) = 0$ 算子约束的最优化问题. 称为精确惩罚的这个方法在凸规划的情况下追溯到 Eremin[406] 和 Zangwill[1354], 并且与数值最优化一起已得到很好地发展, 例如, 参见 Bertsekas[111], Burke[188, 189], Polyak[1097] 及其参考文献. 关于在非光滑最优化必要最优性条件中的应用, 这个方法可能首次由 Ioffe[588] 提出, 他建立了定理 5.16; 请比照 Clarke[249, 255] 关于精确惩罚稍微不同的结果, 其中没有特别研究算子约束. 关于精确惩罚技术在约束最优化、变分法和最优控制问题的必要最优性条件中的各种应用, 建议读者参阅 Demyanov 近期的书 [318] 及其参考文献.

在定理 5.16 中实施的主要概念是在一点处的正则性 (在定义 5.15 中称为弱度量正则性), 它由 Ioffe 在文献 [587] 中引入. 这个概念密切关联于 Dontchev 与 Rockafellar[366] 术语下的次正则性, 一般来说, 不同于全书中使用的在一点附近的度量正则性这个基本概念. 弱度量正则性不是鲁棒的, 从而不能得到充分地刻画以及类似基本度量正则性的分析法则/保持性质. 另一方面, 这个弱度量正则性和相关 (逆) 的平静性概念却方便于各种应用; 参看 5.5.16 小节中更多的评论.

给出具有等式算子约束的 Lipschitz 问题的下次微分最优性条件的定理 5.17 及其具体到广义 Fredholm 型算子约束的推论 5.18 是新的. 与前面提到的 5.5.6 小节中讨论的 Ioffe[598] 和 Ginsburg 与 Ioffe[506] 的结果相比, 新的结果用了比有限余维性质更弱的序列法紧性假设, 并且利用次梯度和法向量这些更小的集合. 但另一方面, 这些结果要求两个空间 X 和 Y 都有 Asplund(一般来说不可分的) 空间结构, 而文献 [598, 506] 中的结果适用于任意 Banach 空间 X, 而空间 Y 上要求有一等价范数使其对偶是严格凸的 (因此 Y 差不多是可分的).

还注意到定理 5.17 中关于算子约束映射 f 的严格 Lipschitz 假设 (弱于文献 [506, 595, 598] 中对 f 的严格预可导假设) 可松弛到 f 的几乎局部 Lipschitz 连续

性, 但此时定理 5.17 的乘子规则中的标量化 $\partial\langle y^*, f\rangle(\bar{x})$ 的基本次微分应以 (更大的) 基本 (法锥) 上导数 $D_N^* f(\bar{x})(g^*)$ 代替. 注意到, 这样的上导数形式即使在有限维空间中也没有 Clarke 结构的任何对应物.

5.5.9 有限多泛函约束下的必要最优性条件

5.1.3 小节涉及具有有限多等式、不等式和几何约束的数学规划的更一般形式 (5.23). 这样的约束最优化问题是 5.1.2 小节考虑的具有算子约束的最优化问题的特殊形式, 而具体形式 (5.23) 允许建立关于必要最优性条件的更多种方法和结果.

定理 5.19 中上次微分条件部分是新的, 部分取自 Mordukhovich[925]. 注意到定理 5.19(i) 的最优性条件不但对如 5.2.1 小节的费用函数 φ_0, 而且对描述 (5.23) 中的不等式约束的函数 φ_i $(i = 1, \cdots, m)$ 使用了 Fréchet 上次梯度. 然而这样一来 Fréchet 次梯度的光滑变分描述在证明中就是至关重要的, 这要求所论空间 X 的特殊“光滑阻尼”结构.

5.1.3 小节的随后结果处理非可微规划问题 (5.23) 的下次微分条件, 不仅包括利用费用函数和不等式约束的下次梯度的必要最优条件, 而且包括由相应上图的广义法锥表达的条件.

该小节从定理 5.21(i) 和 (ii) 中这样的几何条件开始, 在一般 Asplund 空间框架下给出关于 (5.23) 的必要最优性条件的近似/模糊和确切/点基两种形式, 其中所涉及的函数无 Lipschitz 假设, 这是由极点原理相应形式直接得到的. 这些最一般形式的结果首次在文献 [922] 中给出, 但实际上这些结果以及所用方法可追溯（有时作为最优控制中的横截条件）到 Mordukhovich 原创性的 [887, 889, 892] 和 Kruger 与 Mordukhovich[717 – 719], 其中这种类型的必要最优性条件为有限维和 Fréchet 光滑空间中问题 (5.23) 的各种特殊情形而建立.

如定理 5.21(iii) 给出的点基条件的次微分形式, 其中在相应函数的局部 Lipschitz 连续性的条件下用其基本次梯度取代了其上图的基本法向量. 这些结果在有限维、Fréchet 光滑和 Asplund 空间框架下也可在前面提到的文章中找到. 注意到在文献 [706, 707] 中 Kruger 得到了一个关于以包含形式 $f(x) \in \Theta$ 给出的具有无限多不等式约束的问题这些结果的推广, 这里 f 是单值 Lipschitz 映射, 同时 $\Theta \subset Y$ 是 Fréchet 光滑空间的上图 Lipschitz 子集. 后面的这个条件当 Θ 是凸集时简化为 $\operatorname{int}\Theta \neq \varnothing$; 这是其中“无限多不等式”名称的由来. 这样的“不等式型”结果可以从定理 5.8(iv) 在 Asplund 空间中在更弱的 SNC 假设下而得.

下面讨论问题 (5.23) 中等式约束

$$\varphi_i(x) = 0, \quad i = m+1, \cdots, m+r$$

的处理, 其中与 5.1.3 小节中上和下次微分条件处理方法相同, 而大大地不同于对

不等式约束和所考虑的费用函数的处理方法. 当 φ_i $(i=m+1,\cdots,m+r)$ 是局部 Lipschitz 的, 等式约束可通过存在任意 (无符号) Lagrange 乘子和 $\lambda_{m+1}\varphi_{m+1}+\cdots+\lambda_{m+r}\varphi_{m+r}$ 的基本次微分

$$\partial\left(\sum_{i=m+1}^{m+r}\lambda_i\varphi_i\right)(\bar{x}),\quad (\lambda_{m+1},\cdots,\lambda_{m+r})\in\mathbb{R}^r, \tag{5.120}$$

这个 "压缩项" 在必要最优性条件中反映出来; 特别地, 参见定理 5.19 中的条件 (5.27). 由于 λ_i 在 (5.120) 中不一定是非负的, 而基本次微分 ∂ 一方面满足和法则, 另一方面又是一单边结构, 因此可用更大的 "集合和"

$$\sum_{i=m+1}^{m+r}\lambda_i\partial^0\varphi_i(\bar{x}),\quad (\lambda_{m+1},\cdots,\lambda_{m+r})\in\mathbb{R}^r$$

来代替 (5.120), 这个和由各个等式约束函数 φ_i 的对称次微分 $\partial^0\varphi_i(\bar{x})=\partial\varphi_i(\bar{x})\cup\partial^+\varphi_i(\bar{x})$ 描述, 而不仅仅是基本次微分 $\partial\varphi_i(\bar{x})$; 请比照推论 5.20. 但这里更喜欢用更确切但不太常见的次微分来表示 (5.120), 此时所有的乘子都是非负的:

$$\sum_{i=m+1}^{m+r}\lambda_i\Big[\partial\varphi_i(\bar{x})\cup\partial(-\varphi_i)(\bar{x})\Big],\quad 其中\quad \lambda_i\geqslant 0,\quad i=m+1,\cdots,m+r.$$

参看定理 5.21(iii) 及其证明, 其中证明中包含当约束函数 φ_i $(i=m+1,\cdots,m+r)$ 在 $\bar{x}$ 附近是局部 Lipschitz 时, 利用包含关系 (5.32), 上面最后一个表达式可由定理 5.21(ii) 中的几何条件 $(x_i^*,-\lambda_i)\in N((\bar{x},0);\operatorname{gph}\varphi_i)$ 得到. 这明显地将定理 5.21(iii) 中的下次微分最优性条件与非可微规划中 Lagrange 乘子法则的其他版本区别开, 特别地, 与 Clarke[249] 和 Warga[1319] 利用其双边次微分结构同等地处理不等式和等式约束的结果相区别, 更多讨论和说明性的例子参见注 5.22.

5.5.10 Lagrange 原理

5.1.3 小节的下一课题, 关联于 (5.23) 型的约束优化问题的下次微分条件, 涉及所谓的 Lagrange 原理的非光滑推广. 这个名称由 Tikhomirov 提出 (特别地, 参见他与 Ioffe[618], Alekseev 与 Fomin[7] 和 Brinkhuis[178] 等书), 他注意到产生在数学和应用科学 (非线性规划、变分法、最优控制、逼近理论、不等式、经典力学、天文学、光学等) 各种领域中的许多极值问题的必要最优性条件可用下列程序得到: 用公式 (5.35) 定义 Lagrange 函数 $L(x,\lambda_0,\cdots,\lambda_{m+r})$, 其中乘子 $(\lambda_0,\cdots,\lambda_{m+r})$ 对应于费用函数和所有泛函 (等式和不等式型) 约束, 然后考虑最小化受约束于剩下的几何约束的 Lagrange 函数. Lagrange 原理表明, 依据 Lagrange[737] 的主要思想, 原

来的约束问题的必要最优性条件可由极小化只受几何约束的具有某非平凡 Lagrange 乘子的 Lagrange 函数的必要最优性条件而得到, (即, 如果在原来的问题中没有几何约束, 那么就是完全无约束的).

当然, 对每一类型最优化问题 Lagrange 原理的有效性都应被证明. Ioffe 和 Tikhomirov 在他们的书 [618](最初以俄文于 1974 年出版) 中对具有"光滑凸"结构的极值问题证明了这一点. 这些极值问题包括涉及光滑动态、不等式型的状态约束和关于控制函数的一般几何约束的问题.

Lagrange 原理的第一个非光滑版本由 Hiriart-Urruty 对 (5.23) 型 Lipschitz 问题利用 Clarke 广义梯度和法锥在文献 [571] 中得到. 对算子约束问题, 进一步的结果由 Ioffe[595] 利用 Clarke 结构建立 (参见 5.5.6 小节), 接着由 Kruger[707], Mordukhov- ich[897, 901] 和 Ginsburg 与 Ioffe[506] 利用非凸次微分结构建立.

引理 5.23 和定理 5.24 的结果是新的; 一些特殊情形和推论可在文献 [707, 708, 897, 901] 中找到. 关于"抽象极大值原理"的推论 5.25 揭示了在变分分析中得到充分认识的事实: 极大值型最优性条件通过凸集法锥的极值结构关联于几何约束的凸性. 在此方向上, 值得注意的是, 连续时间系统最优控制中的极大值原理一般不要求任何明确的凸性假设, 这源于这些系统固有的某种"隐藏的凸性"; 更多细节和讨论参见第 6 章.

5.5.11 混合乘子法则

众所周知, 在最优化理论中等式和不等式约束具有本质上不同的性质, 因此它们应该被区别对待. 如上所见, 非光滑最优化问题中的等式和不等式约束可用相应的必要最优性条件中的不同次梯度集合来区分. 注意到, 费用函数在必要条件中通常作为描述不等式约束的条件处理.

定理 5.26 给出了问题 (5.23) 另一类型的下次微分最优性条件, 明确地区分这个问题中的等式和不等式约束. 这个定理的本质部分由 Mordukhovich[897, 901] 在有限维空间中最先建立, 给出了 Lagrange 乘子法则的混合次微分推广. 事实上, 当基本鲁棒次微分 (严格导数的推广) 被用于 Lagrange 函数必要最优性条件的等式约束时, 经典导数的非鲁棒推广被用于不等式约束和目标函数.

在定理 5.26 中应用的"上凸逼近"的概念和生成的"p- 次微分"结构 (5.47) 归功于 Pshenichnyi[1108, 1109]. 注意到这些概念一般来说被非唯一性和非构造性地定义. 可能的上凸逼近之一是 Clarke 方向导数, 它通常不是最好的逼近, 正如前面提及的 Pshenichnyi 的工作说明的那样. 另一方面, 易证任何 Gâteaux 可微函数通过它的 Gâteaux 导数给出最好的上凸逼近 (参见定理 5.26 之后的讨论), 从而给出了 Lagrange 乘子法则的一个版本, 一般来说它独立于 5.3.1 小节前面的必要最优

性条件.

5.5.12 非 Lipschitz 数据问题的必要条件

由上面给出的结果和讨论可看到, 非可微规划问题 (5.23) 广义 Lagrange 乘子型必要最优性条件的所有下次微分版本都是在描述目标和泛函约束函数 φ_i $(i = 0, \cdots, m+r)$ 的局部 Lipschitz 假设下得到的. 泛函数据只有可微性但无严格/连续可微性 (即一般不具局部 Lipschitz 连续性) 假设下也有关于 Lagrange 乘子经典微分形式的结果; 这在前面有部分地讨论过; 参见定理 5.7(i) , 定理 5.11(i) 以及 Halkin[543], Ioffe[595] 和 Ye[1340, 1341] 的文章. 与此同时, 定理 5.21(ii) 在没有 Lipschitz 假设下给出了问题 (5.23) 在参考极小值点 $\bar{x}$ 处的必要条件, 但不具有传统次微分形式: 它涉及图和上图的基本法锥, 即, 接下来不但涉及费用和约束函数的基本次梯度, 而且涉及奇异次梯度.

定理 5.28 中非 Lipschitz 问题近似/模糊形式的另一下次微分最优性条件由 Borwein, Treiman 与 Zhu[158] 在自反空间中首次得到, 其中所论空间的自反性在证明中是举足轻重的. 也可参见 Borwein 与 Zhu 的文章 [163, 164]. 这种弱模糊最优性条件的 Asplund 空间版本由 Mordukhovich 与 B. Wang[962] 和 Ngai 与 Théra[1009] 利用不同的证明方法独立地得到. 本书所给的证明取自文献 [962]. 对在相对于某生成族具有光滑重赋范的 Banach 空间中非可微规划问题, 这种类型的结果还由 Zhu 在文献 [1737] 中得到.

5.5.13 次优性条件

5.1 节的最后一小节主要研究约束优化问题的次优性条件. 这是指这样一类结果, 由之建立几乎最优解的几乎必要最优性条件, 其中“几乎”意思是“达到任意 $\varepsilon > 0$”这个程度.

从实际应用的观点以及基于必要条件证明数值算法合理的观点来看, 必要最优性和次优性条件之间似乎没有太多明显的差别. 次优性与必要最优性条件相比, 主要的好处在于处理次优性允许我们避免最优解存在性的这个主要的困难, 即最优解可能不存在 (尤其在无限维中), 或者可能难于验证它们的存在性.

次优性条件的重要性在有重大影响的 Young[1349, 1350] 和 McShane[861, 862] 之后在经典的变分法中得到很好地认识. 这些基础结果主要目的不但在于构造可用原始问题次优解逼近有在“广义曲线”类中的最优解的变分问题, 而且在于建立“广义曲线”的必要最优性条件, 这恰好给出了“普通曲线”极小化序列的次优性条件.

由 Gamkrelidge[495 – 497] 和 Warga[1313 – 1315] 在最优控制理论中继续发展了这个方向, 他们利用稍微不同但是等价的凸化程序独立地构造了原始控制问题

的一个恰当松弛 (该术语由 Warga 首次使用), 并且通过必要最优性条件及其广义/松弛对应物的近似最终得到原始控制的极小化序列的次优性条件. 关于这些方法和结果与经典的变分法之间关系的讨论, 也可参见 Ioffe 与 Tikhomirov[617], McShane[863] 和 Young[1351] 的文章. 动态最优化和各种控制问题的次优性条件在没有利用任何松弛程序的条件下被其他研究人员在下列文献中随后得到: Gabasov, Kirillova 与 Mordukhovich [488], Gavrilov 与 Sumin[500], Medhin[867], Mordukhovich [901], Moussaoui 与 Seeger[987], Plotnikov 与 Sumin[1084], Seeger[1199], Sumin[1233, 1234], Zhou[1367–1369] 等.

对 Banach 空间中的一般 (不一定是动态) 最优化问题而言, 第一个次优性条件由 Ekeland 利用他强有力的变分原理在 [369, 397, 399] 中建立. 正如 [397] 中提到的那样, 这样的次优性课题是建立 Ekeland 变分原理的最初动力之一. 对于具有类似 (5.23) 中等式和不等式约束数学规划问题, Ekeland 在初始数据的光滑性假设下在文献 [397] 中得到了 ε- 乘子型规范次优性条件, 其中用了所有约束函数的梯度的线性无关性, 这比 Mangasarian 与 Fromovitz 规范条件更强.

基于 Ekeland 变分原理和合适扰动问题中的必要最优性条件, 对各类非光滑最优化问题下次微分次优化条件能够以规范和非规范两种形式建立, 其中主要通过利用 Clarke 广义微分结构. 读者可在下面的研究结果及参考文献中找到这个方向若干结果和应用: Attouch 与 Wets[47], Auslender 与 Teboulle[60], Bustos[207], Dutta[374], Gupta, Shiraishi 与 Yokoyama[526], Hamel[546], Kusraev 与 Kutateladze[733], Loridan[811], Loridan 与 Morgan[812], Moussaoui 与 Seeger[986].

5.1.4 小节给出的结果取自 Mordukhovich 与 B. Wang 的文章 [962], 该文章基于定理 2.28 中下次微分变分原理的应用和广义微分分析法则的合适技巧. 这里区别两种主要类型次优性条件: 定理 5.29 中的弱条件和定理 5.30 及其推论以规范和非规范两种形式给出的强条件.

实际上, 定理 5.29 中的弱次优性条件在 Asplund 空间情形下对初始数据不要求任何假设 (除了费用和不等式约束函数的下半连续性, 描述等式约束的函数的连续性和几何约束的闭性这些最少的局部假设), 但是所得结果涉及原点的弱 * 邻域 $V^* \subset X^*$, 它利用在参考极小值点附近的 Fréchet 法锥和次梯度给出了 Lagrange 乘子法则的一弱化模糊版本.

相反, 定理 5.30 中的规范形式和推论 5.32 中的非规范形式的强次优性条件建立了近似 Lagrange 乘子规则的一个更合适的强版本, 用小的对偶球 $\varepsilon\mathbb{B}^*$ 代替定理 5.29 中的弱 * 邻域 V^*, 在额外的 Lipschitz 和 SNC 假设下利用基本法锥和次梯度表示. 特别注意推论 5.31 的结果, 它在经典 Mangasarian-Fromovitz 约束规范的条件下给出了光滑非线性规划问题的强次优性条件.

5.5.14 具有均衡约束的数学规划

5.2 节研究的一般类约束最优化问题现称为具有均衡约束的数学规划 (MPEC). 这个名称出现在 Luo, Pang 与 Ralph 的书 [820] 中, 该书包含有限维中这个引人注目的数学规划类的各种定性和数值结果及实际应用. 研究 MPEC 和相关最优化的另一 (非光滑) 方法由 Outrata, Kočvara 与 Zowe 在书 [1031] 中建立.

从历史的角度看, MPEC 可追溯到 20 世纪 30 年代关于分层最优化问题 (现称为 Stackelberg 对策) 的经济学文献中; 关于最初的动因和应用参见 von Stackelberg 的书 [1222], 现代的重新描述可参见 Leitmann 的文章 [758]. 这类分层问题密切关联于双层规划, 其中的焦点是以这样一种方式相互关联的两个层次的数学规划, 其中下层参数问题的最优解集是上层问题的可行解集. 关于双层规划更多的结果、参考和讨论读者可在 Dempe 的书 [316], 他的全面的 (直到 2003) 注释参考书目 [317] 和 Dutta 与 Dempe 的近期文章 [377] 中找到.

在分层最优化中, 经常不仅要考虑下层问题的最优解, 而且要考虑所谓的 KKT 点组成的更大集合, 这包括最优 (或稳定) 解以及一阶最优性条件中相应 Lagrange 乘子的集合. 由此上层问题可行解集的描述涉及具有不等式约束的数学规划的经典互补松弛条件. 由于它们自身的原因, 这种类型的条件具有重大意义; 它们在数学规划中充分发展的互补理论中已被研究多年; 例如, 关于有限维空间中各类互补问题的广泛研究, 参见 Cottle, Pang 与 Stone 的书 [294] 和 Facchinei 与 Pang 不久前的两卷专著. 分层最优化 (下层) 框架下的互补条件引起了具有互补约束的数学规划 (MPCC) 的研究, 这是 MPEC 理论和应用中最重要的部分之一.

另一方面, 还有重要 MPEC 类, 它们的可行解由比互补性更一般的条件给出; 特别地, 由参数变分不等式定义的可行解; 例如, 参见前面提到的书 [424]. 众所周知, 描述 MPEC 可行解最自然和方便的方案, 涵盖前面提到的情形和其他引人注目的非传统数学规划类, 是参数广义方程 (5.53) 的 Robinson 框架. 这种方法被证明不但适用于定义下层最优化、互补和变分不等式问题的最优解/KKT 点的集合, 而且适用于产生在经济学、力学及其他应用科学中的各类均衡. 这也完全解释了名称“均衡约束”广泛传播在与最优化相关的文献中的原因.

MPCC 和 MPEC 的标志性性质是其内蕴的非光滑性, 即使对具有光滑初始数据的问题亦如此. 这样的非光滑性有时是隐性的, 然而在理论和数值算法中仍起着至关重要的作用. 因此一点也不令人惊奇的是, 非光滑分析和广义微分的方法恰好在建立 MPCC 和 MPEC 的各种理论和计算方面提供了重要的工具, 尤其在关联必要最优性条件、灵敏性和稳定性分析、收敛率和误差估计方面.

广义微分的合适概念和相关分析法则的使用导致相应的平稳性概念, 这对 MPEC 的理论和应用来说非常重要. 特别地, B(ouligand)- 平稳性, C(larke)- 平稳

性和 M(ordukhovich)- 平稳性. 最后一个概念近年来已引起较大关注, 这源于一些实际应用 (尤其在力学平衡中) 和它作为 MPEC 的一阶必要最优性条件要求最弱的约束规范. 读者可在下列文献中找到处理 MPEC 平稳性的许多定性和数值结果: Anitescu [20], Facchinei 与 Pang [424], Flegel [454], Flegel 与 Kanzov [455, 456], Flegel, Kanzov 与 Outrata [457], Fukushima 与 Pang [480], Hu 与 Ralph [584], Kočvara, Krŭzik 与 Outrata [689], Kočvara 与 Outrata [690, 691], Outrata [1024-1030], Ralph [1116], Ralph 与 Wright [1117], Scheel 与 Scholtes [1191], Scholtes[1192], Scholtes 与 Stohr[1194], Treiman [1268], Ye[1338, 1339, 1342], Ye 与 Ye [1343], Zhang [1361] 等.

5.5.15 利用基本分析法则的 MPEC 的必要最优性条件

在 5.2.1 小节和 5.2.2 小节建立一般 MPEC 及其特例必要最优性条件的方法是, 首先考虑具有由一般集值映射 $y \in S(x)$ 给出的均衡约束的 (5.52) 型的抽象 MPEC, 然后简化为在 5.1.1 小节研究的只具有几何约束定义在乘积空间中的数学规划, 最后应用关于基本法锥、上导数和次微分结构的广义微分和 SNC 分析法则. 要强调的是, 得到一般 (或更特殊的)MPEC 必要最优性条件的这个方法可以避免 MPEC 研究中利用各种传统的方法简化 MPEC 为通常的数学规划时出现的已被充分认识到的障碍, 此时, 即使在具有光滑数据的简单 MPCC 情形下, 标准的约束条件也可以不满足; 关于更多参考和讨论, 例如, 参见 Ye 的文章 [1338, 1339].

5.2.1 小节和 5.2.2 小节给出的关于一般 MPEC 及其特例的下和上次微分最优性条件几乎都取自 Mordukhovich 的近期文章 [925], 其中有一些是新的. 值得注意的是 (对各类 MPEC 和 MPCC), 下次微分型必要最优性和规范条件以前由 Outrata, Treiman, Ye, Zhang 与他们的合作者利用基本法向量、上导数和次梯度在有限维空间中建立; 参见文献 [457, 689, 690, 691, 816, 1024-1028, 1030, 1268, 1338, 1339, 1342, 1343, 1360, 1361], 在这些文献中, 读者可找到利用初始问题数据表示的许多有效条件以及例子和应用.

关于在 5.2.1 小节中对一般/抽象 MPEC 所得到的必要最优性条件, 注意约束规范中均衡映射 $S(\cdot)$ 的混合上导数与部分 SNC 性质和定理 5.33 与定理 5.34 法紧性假设的关键作用. 由此本书深入地探索了 (在无限维中)MPEC 中内在的基本决策参数空间的乘积结构, 这在很大程度上把这类引人注目的约束最优化问题与具有几何约束的一般数学规划区别开来. 由于这些假设对类 Lipschitz 映射自动成立, 故在有限维和无限维空间中, 所得结果也指出了一类重要的且相当广泛的 MPEC 问题, 它们的一阶规范必要最优性条件总满足; 见推论 5.35.

随后在 5.2.2 小节中对结构化 MPEC 所得的必要最优性条件可看成是 5.2.1 小节中“抽象”MPEC 结果和完善的广义微分和 SNC 分析法则的推论. 进一步, 为

得到灵敏性分析, 大量地应用了 4.4 节得到的参数变分系统的上导数的计算和上估计. 现在灵敏性分析在 MPEC 的必要最优性条件中被充分地应用, 这揭示了这些看似不同课题之间的密切关系. 还注意到由广义变分不等式 (GVI) 控制的 (5.60) 和 (5.63) 型最重要 MPEC 类及其特例的一阶最优性条件中用到了二阶次微分. 然而这并不令人吃惊, 由于 MPEC 约束本身积聚关于下层参数问题的一阶信息; 见上述讨论.

5.5.16 MPEC 最优性条件中的精确惩罚和平静性

5.2.3 小节的结果基于另一方法得到具有由参数变分系统控制的 (5.56) 型均衡约束 MPEC 的必要最优性条件: 除了应用广义微分和 SNC 分析法则之外, 它还包括初始的精确惩罚程序; 请比照 5.1.2 小节中具有等式型算子约束的最优化问题的相应结果. 用这种方法可得 5.2.2 小节建立的由参数变分系统控制的 MPEC 的一些下 (而不是上) 次微分条件的改进.

关于在广义方程约束 (5.96) 下最优化问题的精确惩罚的引理 5.47 由 Ye 与 Ye 在文献 [1343] 中建立, 关于先前的上 Lipschitz 版本也可参见 Zhang [1360]. 注意到它类似于等式约束下最优化问题的精确惩罚结果定理 5.16, 这归功于 Ioffe[588]. 另外, 用于引理 5.47 的定义 5.46 中的平静性条件可看做定义 5.15 中弱度量正则性的一 (逆) 集值对应概念.

在定义 5.46 的框架下, 集值映射的平静性术语由 Rockafellar 与 Wets 在文献 [1165] 提出. 正如 5.2.3 小节在这个定义之后所提到的那样, 在 (5.68) 中, 当 $V = Y$ 时 F 在 $\bar{x} \in \operatorname{dom} F$ 的平静性由 Robinson[1130] 作为集值映射的“上 Lipschitz”性质引入. 在文献 [1132] 中, Robinson 建立了一个关于上 Lipschitz 性质对有限维空间之间的分片多面体集值函数成立的重要事实. 在 $(\bar{x}, \bar{y}) \in \operatorname{gph} F$ 的图局部化版本的平静性 (上 Lipschitz) 性质后来由 Klatte[684] 在不同名称下建立, 接着由 Ye 与 Ye[1343] 在引理 5.47 的背景下独立地建立并有随后 MPEC 中的应用.

另一方面, 最优值函数的平静性, 在与定义 5.46 一致的意义下由 Clarke[249, 255] 作为必要最优性条件的一种约束规范建立 (而由 Rockafellar 提出; 参见文献 [249, P. 172]). 这种约束规范的性质密切关联于由 Dolecki 与 Rolewicz[341] 在精确惩罚的框架下引入的“Φ_1- 次微分”的概念. 关于平静性和相关性质在最优化和变分分析的各个方面中的许多应用, 还建议读者参看文献 Burke [188, 189], Facchinei 与 Pang [424], Henrion 与 Jourani[559], Henrion, Jourani 与 Outrata [560], Henrion 与 Outrata [561, 562], Klatte 与 Kummer [686], Outrata [1027, 1030], Ye [1338, 1339], Zhang[1361, 1362], Zhang 与 Treiman [1363] 和其中的参考文献.

定理 5.48, 定理 5.49 和推论 5.50 的必要最优性条件作为最一般性的结果是新的; 它们的有限维版本和一些具体化由 Outrata[1024, 1027], Ye[1338, 1339], Ye 与

Ye[1343] 和 Zhang[1360] 得到, 并且在一些特殊类 MPEC 中, 尤其在双层规划中, 有各种应用. 关于多面体问题的推论 5.51 和随后的例子取自 Outrata[1027].

5.5.17　多目标最优化和均衡的约束问题

5.3 节主要研究多目标最优化的约束问题, 其中目标函数由一般序关系给出, 特别涵盖了许多不同的均衡. 有大量的文献处理多目标/向量优化和均衡模型的各个方面, 包括最优和均衡解的存在性、最优性条件、数值算法和应用. 关于各种方法、结果和讨论建议读者参看文献 [90, 93, 178, 230, 255, 265, 293, 306, 378, 395, 402, 424, 446, 480, 504, 516, 532, 534, 550, 627, 628, 636, 697, 707, 813, 820, 897, 901, 926, 928, 958, 995, 1000—1002, 1029, 1031, 1040, 1046, 1119, 1134, 1181, 1195, 1214, 1312] 及其中的参考文献. 特别值得注意的是, 上述大多数参考文献并不专门处理经济学模型和相应竞争均衡的概念, 它们将在第 8 章中研究.

5.3 节给出的素材主要包含多目标最优化和均衡问题及其特例的最优解的一般概念. 所得结果的最初目标是得到具有各种约束的某些不寻常类型多目标最优化问题的必要最优性条件, 包括一类新的所谓具有均衡约束的均衡问题 (EPEC), 这类问题在许多实际应用中非常重要. 本小节给出的结果表明, 变分分析和广义微分的高等方法恰好对处理这样的问题是非常有用的, 并且产生强有力的最优性条件, 其中大部分是新的, 或刚刚最近出版. 这里不考虑多目标最优化、均衡的存在性和数值课题, 它们与本书建立的方法很不一样.

本书主要研究多目标最优化和均衡的两种不同方法, 它们明显不同, 即使从解概念的观点来看亦如此. 同时, 利用这些方法所得的约束问题的必要最优性条件一般来说基于极点原理的不同版本.

5.5.18　多目标最优化中的解的概念

定义 5.53 中的广义序最优性的概念可追溯到 Kruger 与 Mordukhovich 早期的工作 (参见文献 [707, 719, 897, 901]); 它由集合系统局部极点的概念直接诱导. 注意到这个抽象的最优性概念对定序集 Θ 不要求任何凸性和/或非空内部的假设; 例如对照 Gamkrelidge[496], Gorokhovik[516], Neustadt[1001, 1002], Rubinov[1181], Singer[1214],Warga[1319] 及其他关于抽象最优性的文献. 定义 5.53 之后讨论的最优性的特殊概念本质上是经典的; 它们大体上可追溯到 Pareto 开创性的工作 [1053]. 注意到研究弱 Pareto 最优解与 (真)Pareto 最优解相比要容易得多; 向量最优化的主要结果实际上都是针对弱 Pareto 解的.

闭偏好关系的定义 5.55 归功于 Mordukhovich, Treiman 与 Zhu[958], 而各种抽象偏好关系已经被长期地研究, 并且被应用到向量最优化中, 尤其是经济学模型中; 例如, 参见 Debreu [310], Mas-Colell [854, 855], Pallaschke 与 Rolewicz [1040], Zhu

[1732] 及其中参考文献. 利用序锥的特殊性质刻画广义 Pareto 序的几乎传递性的命题 5.56 的结果以及表明在有限维空间中字典序可能不具有这种性质的例 5.57 都取自 Eisenhart 在 Zhu 指导下完成的博士论文 [395].

5.5.19 广义序最优性的必要条件

5.3.2 小节给出了一般约束多目标最优化问题及其特例的必要最优性条件, 其中广义序最优性的概念按定义 5.53 的意义来理解. 所得结果基于确切极点原理在引理 5.58 中给出的版本. 它与在 2.2.3 小节建立的版本的主要区别在于它考虑多目标最优化中内在的基础空间的乘积结构. 利用这种方法在无限维空间中可得到关于确切极点原理 (包含 PSNC, 但不包含 SNC 性质) 更多精细条件; 这个方向更进一步的结果参见 Mordukhovich 与 B.Wang[963].

定理 5.59 和它的推论 5.60 是新的; 在有更多限制的假设下一些特殊的结果在 Kruger [707] 和 Mordukhovich [897, 901] 中给出. 定理 5.61 中的上次微分条件也是新的.

很多年来极小极大最优化问题吸引了数学家、应用科学家和实际工作者们的强烈关注, 这是由于其在理论和应用中的特殊重要性. 这样的问题是内蕴地非光滑的, 是利用 (多半特殊的) 非光滑分析的方法研究的最早的非标准的最优化问题之一, 例如, 参见 Clarke[255], Danskin[307], Demyanov 与 Malosemov[319], Dubovitskii 与 Milyutin[370], Ioffe 与 Tikhomirov[618], Krasovskii 与 Subbotin[702], Neustadt[1002], Pshenichnyi[1106], Rockafellar 与 Wets[1165] 以及其中的参考文献.

得到极小极大问题的必要最优性条件的一个 (传统) 方法是利用标量最优化的一般非光滑问题的必要最优性条件, 然后应用计算/估计极大值函数相应的次微分的公式. 用这种方法利用在有限集合上极大值函数的基本次梯度在 3.2.1 小节中的分析法则, 易得到推论 5.63 中的必要最优性条件. 这个结果由 Mordukhovich[892] 在有限维空间中直接应用度量逼近方法首次建立.

用来证明定理 5.62 的方法, 基于广义序最优性的简化, 似乎更适合、更方便处理包含在对偶空间的弱 * 紧子集上的极大化的极小极大问题 (5.83). 定理 5.62 中最一般情形下的结果是新的, 而在极大化下一些紧集的特殊情形以前由 Kruger[706] 和 Mordukhovich[901] 研究过.

5.5.20 极点原理的集值映射推广版本

5.3.3 小节包含极点原理的推广版本, 这对由闭序关系描述的多目标最优化问题的必要最优性条件是需要的. 这样的推广不仅能处理集合的 (线性) 平移, 而且能处理集值映射的非线性变形. 集值映射近似极点原理形式的一个合适结果由定理 5.68 给出, 取自于 Mordukhovich, Treiman 与 Zhu 的文章 [958], 在那里读者可

找到说明关于推广极点系统的定义 5.64 的书中的和额外的例子.

为建立集值映射极点原理的确切版本, 需要移动 (即参数化的) 集合的极限法向量的概念. 一个合适的定义在文献 [958] 中给出, 其中, 在结构 (5.95) 中, 本书置 $\varepsilon = 0$, 在所考虑的 Asplund 空间情形, 这并不限制一般性; 在有限维空间中的结果也可参看 Bellaassali 与 Jourani [93]. 受在集值映射的覆盖性质中应用的促动, (5.96) 中移动集合的法向半连续性的概念更早由 Mordukhovich[894] 引入.

命题 5.70 中法向半连续性的充分条件取自 Mordukhovich [894, 901], 而在 Bellaassali 与 Jourani[93] 中读者可找到这个性质不成立的有趣例子, 其中的集值映射是由在 $\mathbb{R}^2$ 上 Lipschitz 连续效用函数限定序关系的水平集生成的 $S(z) = \operatorname{cl} \mathcal{L}(z)$. 分别受非凸经济和非凸清扫过程的促动, 在有限维中保证一致邻近正则映射法向半连续性的其他充分条件近来已由 Bounkhel 与 Jofré在文献 [171] 中得到, 在 Hilbert 空间中由 Bounkhel 与 Thibault 在文献 [173] 中得到.

正如固定集合的情形需要一定的法紧性来得到无限维中确切/点基形式的结果, 移动集合/集值映射的 SNC 性质的合适推广在定义 5.71 中以"像 SNC"名称表述. 这个性质, 连同它的部分版本以及定义 5.69 中极限法锥结构和相应的次微分与上导数概念, 由 Mordukhovich 与 B. Wang 在文献 [966, 969] 中详细地研究; 参看定义 5.71 之后的一些讨论. 移动集合和映射的推广极限结构也满足分析法则, 类似它们的基本版本, 而基本和推广的 SNC 性质之间的关系更复杂, 依赖于适当定义的一致性.

定理 5.72 中描述的集值映射的确切极点原理由 Mordukhovich, Treiman 与 Zhu[958] 证明. 定理 5.72 中相反的蕴涵关系由定理 2.22(ii) 建立的闭集极点系统的相应结果直接可得.

5.5.21 具有闭序关系的多目标问题的必要条件

5.3.4 小节包含在各种约束 (几何、算子和泛函型) 下多目标最优化问题的必要最优性条件, 其中"多目标极小化"由闭序关系定义. 本小节所得结果基于 5.3.3 小节中的极点原理的推广版本, 并且以近似/模糊和确切/极限形式给出.

对具有几何约束的问题, 定理 5.73(i) 中的模糊最优性条件取自 Mordukhovich, Treiman 与 Zhu [958], 在那里读者也可找到具有有限多等式和不等式型泛函约束的多目标问题的"强"和"弱"模糊形式的必要条件. 在定理 5.73(ii), 推论 5.75 和定理 5.76 中所得的极限最优性条件是新的; 所提到的具有等式和不等式约束问题的部分结果在目标映射 $f\colon X \to Z$ 的值空间 Z 有限维性的假设下在文献 [958] 中得到. 关于具有向量最优性的"广义序"和"闭序"概念的多目标问题所得的相应最优性条件的比较的讨论, 建议读者参看注 5.74.

5.3.4 小节关于多目标对策的素材取自 Mordukhovich, Treiman 与 Zhu [958].

5.5.22 具有均衡约束的均衡问题

5.3.5 小节主要研究具有均衡约束的多目标最优化问题. 这里把这类向量最优化问题看成 5.2 节考虑的 MPEC 在多目标情形的推广. 实际上, 它们包括与 MPEC 相同类型的 (均衡) 约束, 而最优化则是相对于 5.3.1 小节所讨论的一般多目标准则来实施. 正如其中所述, 所考虑的多目标最优化的概念包括均衡的各种概念, 从而这些问题可看做是具有均衡约束的均衡问题 (EPEC).

EPEC 术语在不久前才出现; 它由 Scholtes 在关于互补问题的 2002 国际会议上的报告 [1193] 中首次提出. 实际应用是研究这类多目标最优化问题的原始动因之一; 参看 Hu, Ralph, Ralph, Bardsley 与 Ferris 共同完成的关于无调控电子市场的竞争均衡模型的工作 [585]. 在文献 [585, 1193] 中主要研究 EPEC, 其中领导者 (上层) 和跟随者 (下层) 的行为通过非合作 Nash（或 Cournot-Nash）均衡建模; 请比照文献 [995, 1031]. 关于相关进展也建议读者参看 Fukushima 与 Pang [480] 和 Outrata[1029]. 特别地, 文献 [1029] 包含各种 EPEC 性质的深刻研究, 并且给出了一类非合作 (关于两个层次的) EPEC 的必要最优性条件, 这是通过简化它们为偶联的 MPEC 和利用本书的基本广义微分结构来实现的.

另一种 EPEC 由 Mordukhovich[926, 928] 从上层是多目标最优化、下层是由参数变分系统控制的均衡约束的角度研究. 这样的 EPEC 可自然地当做所有领导者之间互相合作寻找广义 Pareto 型均衡的问题; 它们不能只是简化为 MPEC 系统, 而是需要特殊的考虑. 利用本书给出的基本结构的广义微分分析法则, 由多目标最优化的一般结果, 文献 [926, 928] 在有限维空间中得到了 EPEC 的必要最优性条件. 关于具有变分不等式约束的有限维多目标问题的这种类型的更特殊结果由 Ye 与 Zhu[1345] 得到, 其中上层最优性由某种“正则”序关系定义.

Mordukhovich, Outrata 与 Červinka[940] 近来的工作包含对在上层具有经典的弱 Pareto 最优性, 在下层由互补约束控制的一重要有限维 EPEC 类的 [924, 928] 方法的发展和具体化. 考虑到有限维中互补约束的特殊性质, 文献 [940] 中的必要最优性条件通过初始问题数据构造性给出, 并基于 MPEC 在 Outrata, Kočvara 与 Zowe 的书 [1031] 中建立的隐规划方法, 被用于建立有效数值算法. 而且, 读者在文献 [940] 中可找到所得结果在涉及许多领导者和跟随者分层垄断市场模型中的应用.

5.3.5 小节给出的结果大部分是新的, 推广了 Mordukhovich 先前在文献 [926, 928] 中在有限维空间中所得的最优性条件. 值得注意的是, 无限维情形要明显地复杂, 因为它需要利用广义微分分析法则和合适的 SNC 分析法则的结果通过 EPEC 初始数据来表达必要最优性和规范条件. 注意关于广义序最优性的定理 5.59 和二阶次微分链式法则为得到 5.3.5 小节中 EPEC 必要最优性条件所起的至关重要的

作用.

5.5.23　线性率下的次极性和次优性

5.4 节论及的课题在最优化理论中是非传统的, 传统上必要条件通常 (除凸问题和类似的问题外) 对最优性的标准概念而言不是充分的. 注意到在经典和现代最优化理论的所有分支中, 所有主要的必要最优性条件 (非线性规划中的 Lagrange 乘子和 Karush-Kuhn-Tucker 条件、变分法中的 Euler-Lagrange 方程、最优控制中的 Pontryagin 极大值原理等) 用涉及伴随变量的对偶形式表达. 本书建立的所有广义极性和最优性条件也是这种情形. 同时, 最优性这个概念包括标量和向量的情形当然都用本原空间的术语描述.

一个具有挑战性的问题是找到极性/最优性某种修正的概念, 以便使以前公认概念的已知必要条件在新的框架下变成充要的. 这样的研究由 Kruger[710, 711] 开始, 接着继续于他随后的文章 [713 – 715] 中. 集合极性以及与向量和标量最优化问题相关联的最优性的新概念首次称为“推广的极性/最优性”[710 – 713], 而最近 Kruger[714 – 716] 已开始对它们使用“弱稳定性”的名称. 根据下面所解释的原因, 建议对这些概念使用术语“线性次极性/次优性”(linear subextremality/suboptimality); 请比照 5.4 节的导言部分.

事实上, 新的概念弱于传统的概念, 它关系到所论点附近集合和映射的极值/最优的性质; 因此应用前缀“次”(sub) 对识别这样的性质是有意义的.

这些新概念的另一关键性质是, 与传统的概念相比, 它们涉及极性和最优性的线性率, 类似本书广泛研究的线性开性/覆盖、度量正则性和类 -Lipschitz 性质. 正如所见, 这些重要性质的线性率本质在现代变分分析的框架下 (甚至关于经典的情形) 已被充分地认识到, 是使得可以利用广义微分来建立它们完整刻画的关键.

对 5.4 节研究的 (次) 极性和 (次) 优性概念, 正是这个线性率特征作为驱动力保证了关于传统概念的极性和最优性条件对所考虑的新概念变为充要条件的可能性. 而且, 覆盖/度量正则性/类 -Lipschitz 性质和线性次极性/次优性概念之间有直接的联系, 由本书的两个证明 (例如, 参见定理 5.88 的证明) 和 Kruger 近期文章 [715, 716] 中相应的常数关系揭示.

5.5.24　多目标问题的线性集合次极性和线性次优性

线性集合次极性的定义 5.87 归功于 Kruger[711], 起初被称为“推广的极性”, 接着在文献 [715] 中称为集合系统的“弱极性”. 等价于 (5.106) 形式的线性次极性的充要条件首次由 Kruger[711] 在 Fréchet 光滑空间中提出, 接着在 Asplund 空间情形下在文献 [712, 713] 中被证明. 值得注意的是, 结论 (ii) 的证明类似引理 2.32(ii) 与定理 2.51(i) 的证明. 这些证明分别取自 Mordukhovich 与 Shao [948] 和

Mordukhovich [920]. 利用确切极点原理刻画这个概念的定理 5.89 是新的.

定义 5.91 中多目标问题线性次优性的概念由 Kruger 在文献 [710] 中以“推广的 (f, Ω, Θ)- 极性”的名称引入. 等价于定理 5.92 中 (5.112) 这个概念的模糊刻画首次对 Fréchet 光滑空间在文献 [710] 中提出, 接着对 Asplund 空间在文献 [712] 中被证明. 这一小节关于多目标问题的线性次优性的确切/点基刻画的所有其他结果都是新的.

这些点基刻画涉及利用所论点处的基本法向量、上导数 (混合的和基本的) 和一阶与二阶次梯度. 这允许对这些结构利用完善的广义微分分析法则, 以及无限维中相关的 SNC 分析法则. 要强调的是, 为了用这种方法得到有结构的多目标问题 (包括 EPEC 问题) 的线性次优性的充要条件, 与绝大多数广义微分的应用不同, 不仅需要使用“右”包含型的分析法则结果, 而且需要所有分析层面的等式型结果. 尽管这些分析结果有更多的限制, 但在本书中还是有丰富地发展. 同样地, 也需要利用 SNC 分析法则结果确保在无限维中在各种运算下相应的 SNC 性质之间的等价性.

5.5.25 约束最优化中的线性次极小值

第 5 章的最后一小节涉及关于标量 (增广实值) 函数的最优化问题的线性次极小性的概念. 这个概念由 Kruger[712] 以“几乎极小性”的名称引入, 接着在文献 [713] 中作为“推广极小性”和在文献 [714] 中作为“弱下确界稳定性”被研究. 尽管人们总可以把定义 5.101 中的线性次极小性看成定义 5.91 中描述的对映到广义序空间上的映射的线性次优性的特殊情形, 但是有某些标量情形的特殊性质应该在研究和应用中被考虑. 正如用例 5.102 中的简单函数 (这归功于 Kruger[712]) 所说明的那样, 线性次极小点的性质可能大大地不同于局部极小值点的性质. 另一方面, 在文献 [712] 中已注意到, 对导数为零的光滑函数扰动, 线性次极小点是稳定的, 这与局部极小值点不同. 对光滑函数来说, 这蕴涵着线性次极小性的概念等价于经典的平稳性, 然而对非光滑的情形则不然.

后来在文献 [713] 中 Kruger 所做的另一观察结果是在 Banach 空间上 l.s.c. 函数的一般情形下, 定义 5.101 中的线性次极小性等价于 Kummer[728]“相对于极小化的平稳点”的名称下所引入的概念, 这在定理 5.103(b) 中被描述. 定理 5.103 还通过强斜率的强有力的结构建立了线性次优性的有效描述, 强斜率在发展方程理论中由 De Giorgi, Marina 与 Tosques[312] 引入, 接着被 Azé, Corvellec 与 Lucchetti[70] 和 Ioffe[608] 有效地应用于变分稳定性和度量正则性; 参见 4.5.2 小节的讨论.

定理 5.106(i) 中线性次极小性的模糊次微分准则归功于 Kruger[712], 直接由定理 5.92 得到. 定理 5.106(ii) 是新的. 利用限制函数 $\varphi_\Omega = \varphi + \delta(\cdot; \Omega)$ 的基本次梯度, 它给出了线性次极小性的“压缩”点基刻画, 因此允许在推论 5.107(a)~(c) 的任

一假设下得到方便的“分离型”的准则 (5.114), 这些假设保证次微分和法则作为等式成立; 也可参见该推论后的讨论. 这个结果蕴涵结构化约束极小化问题 (特别地, 对 MPEC) 更多具体准则成立, 这类似 5.4.2 小节基于等式型一阶和二阶分析法则的论证.

最后, 定理 5.108 给出了具有不等式约束的问题线性次极小性新的必要条件. 与线性次优性和次极小性所有以前的结果相比, 它建立了上次微分型条件; 对其他结构化约束优化问题, 也可建立这种类型的条件, 这类似 5.1 节和 5.2 节详细研究的传统最优性的必要条件.

第 6 章　Banach 空间中发展系统的最优控制

下面两章研究最优控制, 它是现代变分分析和广义微分方法最为重要的推动力和最富应用成果的领域之一. 本书使用的基本法向量、次梯度和上导数的概念是由作者在最优控制问题的相关研究中引入并应用的, 这并非偶然. 事实上, 最简单和最初的最优控制问题即是内蕴非光滑的, 既使描述动态和可行轨道约束的函数是光滑的情形亦如此. 与经典的变分法相比, 问题的关键在于, 最优控制问题的一个标志性的特征是控制函数上逐点约束的出现, 这些逐点约束可能 (并且通常) 是由高度不规则的集合来定义的, 例如包含有限多个点的集合. 特别地, 该情形是自动控制中典型的问题. 这些问题为最优控制理论的发展提供了原始动力.

如下展述的主要目标是应用变分分析和广义微分的方法来导出发展型系统的一些最优控制问题的必要最优性条件. 本章涉及 Banach 空间中由常微分方程和包含控制的动态系统; 由泛函微分和偏微分关系控制的具有分布参数系统的控制问题将主要在第 7 章中考虑.

本章的主要焦点放在 Bolza 和 Mayer 型最优控制/动态最优问题上, 它们由具有离散时间和连续时间动态以及端点约束的无穷维发展型包含和控制系统来控制. 除了无穷维变分原理和前面发展的广义微分工具, 还用到了动态最优化和最优控制的特殊技巧. 下面发展的基本方法称为“离散逼近方法”, 由之可利用离散动态控制问题来逼近连续时间控制问题. 连续时间和离散时间控制系统之间的关系是本章的中心议题之一. 无论是从理论还是应用的角度来讲, 由此得到的结果对最优控制的定性和数值方法都很有新意.

6.1　离散时间和连续时间发展型包含的最优控制

本节在适当 (相当广泛) 的 Banach 空间中考虑由微分包含及其有限差分逼近控制的动态/发展型系统的最优控制问题. 所考虑的模型包含了更多由参数化微分方程描述的传统最优控制问题. 研究连续时间控制系统的主要方法是构造适定的离散逼近和建立关于值收敛的变分稳定性以及最优解的适当强收敛. 然后导出由有限差分包含控制的离散时间最优控制问题的必要最优性条件. 后者可化为前一章中考虑过的具有多个几何约束的非动态最优化问题. 另一方面, 它们具有具体的结构特点, 这将在下面探究. 这样一来, 利用第 3 章的广义微分和 SNC 分析法则, 在初始数据上相当一般的假设下, 本节建立了离散逼近问题的确切和模糊形式的必要最

优性条件. 通过应用第 4 章中的 Lipschitz 稳定性的上导数刻画并取极限, 则可在关于速度变量的特定的松弛/凸化假设下得到连续时间系统以 Euler-Lagrange 方程形式给出的“中间局部极小点”, 它介于经典的弱和强极小点之间. 为在恰当的假设下避免提到的这种松弛条件, 下一节发展了一种额外的逼近.

6.1.1 微分包含及其离散逼近

设 X 为 Banach 空间 (以下称为状态空间) 且设 $T := [a, b]$ 为实数时间区间段. 考虑集值映射 $F\colon X \times T \rightrightarrows X$ 并定义由 F 生成的微分/发展包含

$$\dot{x}(t) \in F\big(x(t), t\big), \quad \text{a.e. } t \in [a, b], \tag{6.1}$$

其中 $\dot{x}(t)$ 表示关于时间的导数并且 a.e. (几乎处处) 依惯例表示在 $\mathbb{R}$ 中除去零测度集外均成立. 下面给出本章中用到的微分包含 (6.1) 解的确切定义.

定义 6.1(微分包含的解) 包含 (6.1) 的解指的是映射 $x\colon T \to X$ 对几乎处处的 $t \in T$ 是 Fréchet 可微的, 并且满足 (6.1) 以及 Newton-Leibniz 公式

$$x(t) = x(a) + \int_a^t \dot{x}(s)\mathrm{d}s,$$

对所有 $t \in T$ 成立, 其中的积分是 Bochner 意义下的积分.

众所周知, 对于 $X = \mathbb{R}^n$, $x(t)$ 几乎处处可微并满足 Newton-Leibniz 公式当且仅当在标准意义下在 T 上它是绝对连续的, 即对任给的 $\varepsilon > 0$, 存在 δ, 使得对互不相交的区间 $(t_j, t_{j+1}] \subset T$ 有

$$\sum_{j=1}^{l} \|x(t_{j+1}) - x(t_j)\| \leqslant \varepsilon \quad \Big(\text{若} \sum_{j=1}^{l} |t_{j+1} - t_j| \leqslant \delta\Big).$$

然而, 对于无穷微空间 X, Lipschitz 连续性并不蕴涵几乎处处可微性. 另一方面, 有一个关于 Banach 空间的完全刻画, 其中映射 $x\colon T \to X$ 的绝对连续性等价于几乎处处可微性并且满足 Newton-Leibniz 公式. 此即为具有所谓 Radon-Nikodým 性质 (RNP) 的空间.

定义 6.2(Radon-Nikodým 性质) Banach 空间 X 称为具有 Radon-Nikodým 性质, 如果对每个有限测度空间 (Ξ, Σ, μ) 和每个 μ- 连续的有界变差向量测度 $m\colon \Sigma \to X$, 存在 $g \in L^1(\mu; \Xi)$, 使得

$$m(E) = \int_E g\mathrm{d}\mu \quad \text{对于 } E \in \Sigma \text{ 成立}.$$

这个基本性质在一般的向量测度定理和 Banach 空间几何理论中得到了很好地探究; 对 RNP 及其应用的深入研究推荐读者参阅 Diestel 和 Uhl 的经典著作 [334]

以及 Bourgin [169]. 特别地, 在文献 [334, 217–219 页] 中能找到关于 RNP 等价阐述/刻画的综述以及 RNP 自动成立的具体的 Banach 空间. 值得注意的是, 这些具体的 Banach 空间包含了自反空间和弱紧生成的对偶空间, 因此包含了所有可分对偶空间. 另一方面, 经典空间 $c_0, c, l^\infty, L^1[0,1]$ 和 $L^\infty[0,1]$ 不具有 RNP. 下面用到 RNP 和 Asplund 空间之间的一个很好的关系: 对给定的 Banach 空间 X, 对偶空间 X^* 具有 RNP 当且仅当 X 是 Asplund 空间.

因此, 对具有 RNP 的 Banach 空间 (并且只有对这类空间), 定义 6.1 中解的概念和涉及绝对连续映射的 Carathéodory 解的标准定义是一致的. 一般地, 定义 6.1 中的假设实际上满足这里的要求而不用借助于 Carathéodory 解和 RNP. 然而, 本章的主要 (但不是所有的) 结果都本质地利用了 RNP 和 X 的 Asplund 性质, 只是使用的角度不太一样, 不直接关联到所给的微分包含解的概念.

众所周知, 微分包含 (当然有其独立的意义) 是下述由微分/发展方程控制的并具有控制参数的控制系统的一个有用的推广:

$$\dot{x} = f(x, u, t), \quad u \in U(t), \tag{6.2}$$

其中的控制集 $U(\cdot)$ 也可能通过 $F(x,t) = f(x, U(x,t), t)$ 依赖于状态变量 x. 对于某些情形, 尤其当集合 $F(x,t)$ 是凸的, 微分包含 (6.1) 具有 (6.2) 式的参数表示形式. 但对一般情形, 它们并不能够约化为参数控制系统, 因而需对其专门进行研究. 也注意到, 在 Banach 空间中, (6.2) 形式的常微分方程与抛物和双曲发展型偏微分方程的各种控制问题密切相联, 其中的解以其他某种适当意义来理解; 例如可参见 Fattorini [432], Li 与 Yong [789] 以及注解 6.26 和下面第 7 章中的结果和讨论.

研究微分包含的主要方法涉及以导数的有限差分来替换

$$\dot{x}(t) \approx \frac{x(t+h) - x(t)}{h}, \quad h \to 0.$$

简单起见, 这里只考虑一致 Euler 方法. 为此选取任意自然数 $N \in \mathbb{N}$, 并考虑在 T 上定义的离散格/网

$$T_N := \{\, a, a + h_N, \cdots, b - h_N, b \,\}, \quad h_N = (b-a)/N,$$

具有离散步长 h_N 并且网点为 $t_j := a + jh_N,\ j = 0, \cdots, N$, 其中 $t_0 = a,\ t_N = b$. 那么微分包含 (6.1) 被其一列有限差分/离散逼近所替代:

$$x_N(t_{j+1}) \in x_N(t_j) + h_N F(x_N(t_j), t_j), \quad j = 0, \cdots, N-1. \tag{6.3}$$

给定满足 (6.3) 的一个离散轨道 $x_N(t_j)$, 考虑它在连续时间区间 T 上的分段线性扩张 $x_N(t)$, 即 Euler 折线. 其相应的离散速率在 T 上的分段常数扩张定义为

$$v_N(t) := \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}), \quad j = 0, \cdots, N-1.$$

由 Bochner 积分的原定义可得

$$x_N(t) = x_N(a) + \int_a^t v_N(s)\mathrm{d}s, \quad \text{对 } t \in T \text{ 成立}.$$

下面首要的目标是要证明在合理的假设下, 微分包含 (6.1) 的每一个解能被离散包含 (6.3) 的扩张轨道强逼近. 这里的强逼近理解为经典 Sobolev 空间 $W^{1,2}([a,b];X)$ 依范数拓扑收敛, 其上的范数为

$$\|x(\cdot)\|_{W^{1,2}} := \max_{t\in[a,b]} \|x(t)\| + \left(\int_a^b \|\dot{x}(t)\|^2 \mathrm{d}t\right)^{1/2},$$

其中等式右端的范数是在空间 X 中取的. 注意到, $W^{1,2}([a,b];X)$ 中的收敛蕴涵轨道在 $[a,b]$ 中的 (一致) 收敛及其导数 (或导数的某子列) 对几乎处处 $t \in [a,b]$ 的逐点收敛, 其中的逐点收敛是至关重要的, 尤其对 $F(x,t)$ 取非凸值的情形.

下面阐述对研究所作的假设, 它们不仅在下一个定理中用到, 而且在随后关于以离散逼近来导出的微分包含的结果中也用到. 然而, 这些假设在某些情形可以放宽; 参见以下的注解和讨论. 粗略地讲, 这里假设集值映射 $F\colon X \times [a,b] \rightrightarrows X$ 是紧值的, 关于 x 是局部 Lipschitz 的, 以及在 $[a,b]$ 上关于 t 是几乎处处 Hausdorff 连续的. 更确切地, 以下假设是沿着 (6.1) 的某个轨道 $\bar{x}(\cdot)$ 所作的, 该轨道在下一定理中是任意的, 但以后是所考虑的变分问题的一个参考最优解.

(H1) 存在开集 $U \subset X$ 及正数 m_F 和 ℓ_F, 使得对任意 $t \in [a,b]$ 有 $\bar{x}(t) \in U$; 对所有 $(x,t) \in U \times [a,b]$, 集合 $F(x,t)$ 是非空紧的, 以及包含关系

$$F(x,t) \subset m_F \mathbb{B} \quad \text{对所有 } (x,t) \in U \times [a,b] \text{ 成立}, \tag{6.4}$$

$$F(x_1,t) \subset F(x_2,t) + \ell_F\|x_1 - x_2\|\mathbb{B} \quad \text{对所有 } x_1, x_2 \in U, t \in [a,b] \text{ 成立}. \tag{6.5}$$

(H2) 对几乎处处 $t \in [a,b]$, $F(x,\cdot)$ 关于 $x \in U$ 是一致 Hausdorff 连续的.

注意到, 包含关系 (6.5) 在 X 的非空紧子集的空间上等价于 $F(\cdot,t)$ 关于 Pompieu-Hausdorff 度量 $\mathrm{haus}(\cdot,\cdot)$ 的一致 Lipschitz 连续性

$$\mathrm{haus}(F(x,t),F(u,t)) \leqslant \ell_F\|x-u\|, \quad x,u \in U;$$

参见 1.2.2 小节.

为了有效地处理对几乎处处 $t \in [a,b]$ 的 $F(x,\cdot)$ 的 Hausdorff 连续性, 对 $t \in [a,b]$ 和 $x \in U$, 定义 F 的连续性均模

$$\tau(F;h) := \int_a^b \sigma(F;t,h)\mathrm{d}t, \tag{6.6}$$

其中 $\sigma(F;t,h) := \sup\{\omega(F;x,t,h) | \ x \in U\}$, 并且

$$\omega(F;x,t,h) := \sup\left\{\text{haus}(F(x,t_1),F(x,t_2))\Big|\ t_1,t_2 \in \left[t-\frac{h}{2},t+\frac{h}{2}\right]\cap[a,b]\right\}.$$

以下的结果由定义容易得到.

命题 6.3(连续性均模) 性质 (H2) 成立当且仅当 $h \to 0$ 时, $\tau(F;h) \to 0$.

对于单值映射 $f\colon [a,b] \to X$, 性质 $\tau(F;h) \to 0$ (当 $h \to 0$ 时) 等价于 f 在 $[a,b]$ 上的 Riemann 可积性; 参见 Sendov 和 Popov [1201]. 众所周知, f 在 $[a,b]$ 上 Riemann 可积当且仅当对几乎处处 $t \in [a,b]$, f 是连续的.

下面的强逼近定理在基于离散逼近的更进一步的结果中起到关键的作用.

定理 6.4(离散轨道强的逼近) 在假设 (H1) 和 (H2) 下, 设 $\bar{x}(\cdot)$ 为微分包含 (6.1) 的一个解, 其中 X 是任意的 Banach 空间. 那么存在离散包含 (6.3) 的一列解 $\widehat{x}_N(t_j)$, 使得

$$\widehat{x}_N(a) = \bar{x}(a), \quad \forall N \in \mathbb{N},$$

并且扩张 $\widehat{x}_N(t)$, $a \leqslant t \leqslant b$, 当 $N \to \infty$ 时在空间 $W^{1,2}([a,b];X)$ 中强收敛于 $\bar{x}(t)$.

证明 由涉及 Bochner 积分的定义 6.1, 导数映射 $\dot{x}(\cdot)$ 在 $[a,b]$ 上强可测, 并因此可以在 T 上找到 (如果有必要的话, 重新排列网点 t_j) 一列简单/阶梯映射 $w_N(\cdot)$, 使得对每一个 $j=0,\cdots,N-1$, $w_N(t)$ 在 $[t_j,t_{j+1})$ 上是常数, 并且当 $N \to \infty$ 时, $w_N(\cdot)$ 在 $L^1([a,b];X)$ 中依范数拓扑收敛于 $\dot{x}(\cdot)$. 将该收敛和 (6.1) 及 (6.4) 结合起来, 对所有足够大的 N, 得到

$$\int_a^b \|w_N(t)\|\mathrm{d}t = \sum_{j=0}^{N-1} \|w_N(t_j)\|(t_{j+1}-t_j) \leqslant (m_F+1)(b-a). \tag{6.7}$$

在下面的估计中用到数值序列

$$\xi_N := \int_a^b \|\dot{\bar{x}}(t) - w_N(t)\|\mathrm{d}t \to 0 \quad (N \to \infty).$$

定义离散函数 $u_N(t_j)$ 为

$$u_N(t_{j+1}) = u_N(t_j) + h_N w_N(t_j), \quad j=0,\cdots,N-1, \quad u_N(t_0) := \bar{x}(a)$$

并注意到函数

$$u_N(t) := \bar{x}(a) + \int_a^t w_N(s)\mathrm{d}s, \quad a \leqslant t \leqslant b$$

是 $u_N(t_j)$ 在区间 $[a,b]$ 上的分段线性扩张, 且

$$\|u_N(t) - \bar{x}(t)\| \leqslant \int_a^t \|w_N(s) - \dot{\bar{x}}(s)\|\mathrm{d}s \leqslant \xi_N \quad (t \in [a,b]). \tag{6.8}$$

因此, 只要 N 充分大, 对所有 $t \in [a,b]$, $u_N(t) \in U$.

在 X 的集合上考虑距离函数 $\mathrm{dist}(\cdot;\Omega)$, 可验证 Lipschitz 条件 (6.5) 等价于

$$\mathrm{dist}\big(w;F(x_1,t)\big) \leqslant \mathrm{dist}\big(w;F(x_2,t)\big) + \ell_F\|x_1-x_2\|,$$

其中 $w \in X, x_1, x_2 \in U$ 和 $t \in [a,b]$ 任意; 请比照定理 1.41 的证明. 由 (6.6) 式中 $\tau(F;h)$ 的构造和明显的关系式

$$\mathrm{dist}\big(w;F(x,t_1)\big) \leqslant \mathrm{dist}\big(w;F(x,t_2)\big) + \mathrm{haus}\big(F(x,t_1),F(x,t_2)\big),$$

则有估计

$$\begin{aligned}\zeta_N:&=\sum_{j=0}^{N-1}h_N\mathrm{dist}\Big(w_N(t_j);F\big(u_N(t_j),t_j\big)\Big)\\&=\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\mathrm{dist}\Big(w_N(t_j);F\big(u_N(t_j),t\big)\Big)\mathrm{d}t\\&\leqslant\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\mathrm{dist}\Big(w_N(t_j);F\big(u_N(t),t\big)\Big)\mathrm{d}t+\tau(F;2h_N).\end{aligned}$$

F 的 Lipschitz 性质和 $w_N(\cdot)$ 的构造蕴涵

$$\mathrm{dist}\Big(w_N(t_j);F\big(u_N(t_j),t\big)\Big) \leqslant \mathrm{dist}\Big(w_N(t);F\big(u_N(t),t\big)\Big) + \ell_F(m_F+1)w_N(t_j)(t-t_j)$$

对任意 $t \in [t_j, t_{j+1})$ 成立. 那么

$$\begin{aligned}\mathrm{dist}\Big(w_N(t);F\big(u_N(t),t\big)\Big)&\leqslant\mathrm{dist}\Big(w_N(t);F\big(\bar{x}(t),t\big)\Big)+\ell_F\|u_N(t)-\bar{x}(t)\|\\&\leqslant\|w_N(t)-\dot{\bar{x}}(t)\|+\ell_F\xi_N,\quad \text{a.e. } t\in[a,b].\end{aligned}$$

进一步, 由 (6.7) 和 (6.8), 得到估计

$$\zeta_N \leqslant \gamma_N := (1+\ell_F(b-a))\xi_N + \ell_F h_N^2(b-a)(m_F+1)/2 + \tau(F;2h_N). \tag{6.9}$$

注意到, 上面构造的函数 $u_N(t_j)$ 并不是离散包含 (6.3) 的轨道, 因为 $w_N(t_j) \in F(u_N(t_j),t_j)$ 并不成立. 现在利用 $w_N(t_j)$ 来构造 (6.3) 的靠近 $u_N(t_j)$ 且具有定理所述收敛性质的实际轨道.

利用下面的邻近算法来递归地定义 $\widehat{x}_N(t_j)$, 由 F 的紧值假设, 这是可以实现的:

$$\begin{cases}\widehat{x}_N(t_0)=\bar{x}(a),\quad \widehat{x}_N(t_{j+1})=\widehat{x}_N(t_j)+h_N v_N(t_j),\quad j=0,\cdots,N-1,\\ \text{其中}\quad v_N(t_j)\in F(\widehat{x}_N(t_j),t_j),\\ \text{并且}\quad \|v_N(t_j)-w_N(t_j)\|=\mathrm{dist}\Big(w_N(t_j);F(\widehat{x}_N(t_j),t_j)\Big).\end{cases} \tag{6.10}$$

首先证明只要 N 充分大, 算法 (6.10) 将 $\widehat{x}_N(t_j)$ 保持在假设 (H1) 中的邻域 U 内. 实际上, 对所有 $t \in [a,b]$, 考虑满足 $\bar{x}(t) + \eta_N \mathbb{B} \subset U$ 的任意数 $N \in \mathbb{N}$, 其中

$$\eta_N := \gamma_N \exp(\ell_F(b-a)) + \xi_N,$$

这里的 ξ_N 和 γ_N 在上面已定义. 当 $N \to \infty$ 时, 有 $\eta_N \to 0$, 因为由 ξ_N 的构造有 $\xi_N \to 0$ 以及 $\gamma_N \to 0$ (因为由命题 6.3, 假设 (H2) 等价于 $\tau(F;h_N) \to 0$). 由归纳法, 假设 $\widehat{x}_N(t_j) \in U$ 对所有 $i = 0, \cdots, j$ 成立, 下面证明这对 $i = j+1$ 也成立. 利用 (6.5), (6.9) 和 (6.10), 有

$$\begin{aligned}
\|\widehat{x}_N(t_{j+1}) - u_N(t_{j+1})\| &\leqslant \|\widehat{x}_N(t_j) - u_N(t_j)\| + h_N\|v_N(t_j) - w_N(t_j)\| \\
&\leqslant \|\widehat{x}(t_j) - u_N(t_j)\| + h_N \operatorname{dist}\Big(w_N(t_j); F(u_N(t_j), t_j)\Big) \\
&\quad + \ell_F h_N \|\widehat{x}_N(t_j) - u_N(t_j)\| \leqslant \cdots \\
&\leqslant h_N \sum_{i=0}^{j} (1+\ell_F h_N)^{j-i} \operatorname{dist}\Big(w_N(t_i); F(u_N(t_i), t_i)\Big) \\
&\leqslant h_N \exp[\ell_F(b-a)] \sum_{i=0}^{j} \operatorname{dist}\Big(w_N(t_i); F(u_N(t_i), t_i)\Big) \\
&\leqslant \gamma_N \exp(\ell_F(b-a)).
\end{aligned}$$

由 (6.8), 上面的关系式蕴涵

$$\|\widehat{x}_N(t_{j+1}) - \bar{x}(t_{j+1})\| \leqslant \gamma_N \exp(\ell_F(b-a)) + \xi_N =: \eta_N, \tag{6.11}$$

这就证明了 $\widehat{x}_N(t_j) \in U$ 对任意 $j = 0, \cdots, N$ 成立. 由此, 由前面的讨论有

$$\sum_{j=0}^{N} \|\widehat{x}_N(t_j) - u_N(t_j)\| \leqslant (b-a)\exp(\ell_F(b-a)) \sum_{j=0}^{N-1} \operatorname{dist}\Big(w_N(t_j); F(u_N(t_j), t_j)\Big).$$

现在来估计

$$\vartheta_N := \int_a^b \|\dot{\widehat{x}}_N(t) - w_N(t)\| \mathrm{d}t \quad (N \to \infty).$$

利用上面最后的估计式和 (6.9) 及 (6.11), 有

$$\begin{aligned}
\vartheta_N &= \sum_{j=0}^{N-1} h_N \|\dot{\widehat{x}}_N(t_j) - w_N(t_j)\| = \sum_{j=0}^{N-1} h_N \operatorname{dist}\Big(w_N(t_j); F(\widehat{x}_N(t_j), t_j)\Big) \\
&\leqslant \sum_{j=0}^{N-1} h_N \operatorname{dist}\Big(w_N(t_j); F(u_N(t_j), t_j)\Big) + \ell_F \sum_{j=0}^{N-1} h_N \|\widehat{x}_N(t_j) - u_N(t_j)\| \\
&\leqslant \gamma_N(1 + \ell_F(b-a)\exp(\ell_F(b-a))).
\end{aligned}$$

因此, 最后得到

$$\begin{aligned}&\int_a^b \|\dot{\widehat{x}}_N(t)-\dot{\bar{x}}(t)\|\mathrm{d}t\\ \leqslant&\int_a^b \|\dot{\widehat{x}}_N(t)-w_N(t)\|\mathrm{d}t+\int_a^b \|w_N(t)-\dot{\bar{x}}(t)\|\mathrm{d}t\\ \leqslant&\gamma_N(1+\ell_F(b-a)\exp(\ell_F(b-a)))+\xi_N:=\alpha_N.\end{aligned}\tag{6.12}$$

因为当 $N\to\infty$ 时 $\alpha_N\to 0$, 由对 $\widehat{x}_N(t)$ 和 $\bar{x}(t)$ 的 Newton-Leibniz 公式以及有界性假设 (6.4), 这显然蕴涵着 $W^{1,2}([a,b];X)$ 中所要的依范数收敛 $\widehat{x}_N(\cdot)\to\bar{x}(\cdot)$. △

注 6.5(离散逼近的数值有效性)　由 (6.12) 式和 Newton-Leibniz 公式可得

$$\|\widehat{x}_N(t)-\bar{x}(t)\|\leqslant\int_a^b \|\dot{\widehat{x}}_N(s)-\dot{\bar{x}}(s)\|\mathrm{d}s\leqslant\alpha_N\quad \text{对所有 } t\in[a,b] \text{ 成立}.$$

因此, 定理 6.4 中离散逼近的误差估计和数值有效性依赖于 (6.6) 式中连续性均模 $\tau(F;h)$ 的值和定理 6.4 证明中所定义的逼近量 ξ_N. 令

$$v(F):=\sup_k\left\{\sum_{i=1}^{k-1}\sup_x\Big[\mathrm{haus}\big(F(x,t_{i+1}),F(x,t_i)\big),x\in U\Big],a\leqslant t_1\leqslant\cdots\leqslant t_k\leqslant b\right\},$$

不难验证

$$\tau(F;h)\leqslant v(F)h=O(h),$$

其中只要 $F(x,\cdot)$ 具有有界变差 $v(F)<\infty$ 关于 $x\in U$ 是一致的; 参阅 Dontchev 和 Farkhi [354]. 更进一步, 如果 $\dot{\bar{x}}(\cdot)$ 在 $[a,b]$ 上 Riemann 可积, 对 $t\in[t_j,t_j+h_N)$, 取 $w_N(t)=\dot{\widehat{x}}_N(t)=\dot{\bar{x}}(t_j)$, 则有估计

$$\xi_N\leqslant 2\tau(\dot{\bar{x}};h_N).$$

注 6.6(单边 Lipschitz 微分包含的离散逼近)　定理 6.4 中, F 上的 Lipschitz 连续性和紧值的假设, 在所论的状态空间 X 上的额外要求下可放宽. 特别地, 由 Donchev 和 Mordukhovich[346] 在 Hilbert 空间情形得到了上述定理中一些相应的 $C([a,b];X)$ 逼近和 $W^{1,2}([a,b];X)$ 逼近的结果, 其中的经典 Lipschitz 连续性替换为如下的 F 关于 x 的单边 Lipschitz 性质: 存在常数 $\ell\in\mathbb{R}$ (无需是正的), 使得

$$\sigma\big(x_1-x_2;F(x_1,t)-F(x_2,t)\big)\leqslant\ell\|x_1-x_2\|^2,\quad\forall x_1,x_2\in U,\quad\text{a.e. } t\in[a,b],$$

其中 $\sigma(x;Q):=\sup_{q\in Q}\langle x,q\rangle$ 表示 $Q\subset X$ 的支撑函数. 此外, 映射 $F(\cdot,t)$ 的紧值假设可换为有界集上的有界性; 更多的细节和讨论请参阅所提到的文章.

6.1.2 微分包含的 Bolza 问题与松弛稳定性

本小节开始考虑 Banach 空间中微分包含解 (在定义 6.1 意义下) 上的动态最优化问题: 在微分包含 (6.1) 的轨道 $x\colon [a,b] \to X$ 上考虑如下 Bolza 泛函的极小值

$$J[x] := \varphi(x(a), x(b)) + \int_a^b \vartheta(x(t), \dot{x}(t), t)\mathrm{d}t \tag{6.13}$$

使得 $\vartheta(x(t), \dot{x}(t), t)$ 在固定时间区间 $T := [a,b]$ 上是 Bochner 可积的并且有端点约束

$$(x(a), x(b)) \in \Omega \subset X^2. \tag{6.14}$$

该问题记为 (P) 并称之为微分包含的 (广义) Bolza 问题. 这里使用轨线这个术语表示 (6.1) 满足 $J[x] < \infty$ 的任何解 $x = x(\cdot)$, 并使用可行轨线这个术语表示满足端点约束 (6.14) 的轨线 $x(\cdot)$. 因为动态 (6.1) 和端点约束 (6.14) 都具体给出, 可以假设费用泛函 (6.13) 中的函数 φ 和 ϑ 都取有限值.

问题 (P) 涵盖了有限维和无穷维空间中广泛的动态最优化问题. 特别地, 它包涵了可能具有闭环控制集合 $U(x,t)$ 的具有参数控制系统 (6.2) 的最优控制的标准和非标准模型. 也注意到, 具有自由时间 (无固定时间区间), 具有在 $(x, \dot{x})$ 上的积分约束和具有其他类型的状态约束问题可归结为形式 (P). 为了导出问题 (P) 的最优解的必要条件, 它不仅适用于整体极小点而且也适用于局部极小点. 首先介绍局部极小点的适当概念, 基本的概念如下.

定义 6.7(中间局部极小点) 可行轨线 $\bar{x}$ 称为 (P) 的秩为 $p \in [1,\infty)$ 的中间局部极小点 (i.l.m.), 如果存在 $\varepsilon > 0$ 和 $\alpha \geqslant 0$, 使得对任何满足

$$\|x(t) - \bar{x}(t)\| < \varepsilon \quad (\forall t \in [a,b]), \tag{6.15}$$

$$\alpha \int_a^b \|\dot{x}(t) - \dot{\bar{x}}(t)\|^p \mathrm{d}t < \varepsilon \tag{6.16}$$

的 (P) 的可行轨线 $x(t)$, 有 $J[\bar{x}] \leqslant J[x]$.

关系式 (6.15) 和 (6.16) 实际上意味着考虑 Sobolev 空间 $W^{1,p}([a,b];X)$ 中 $\bar{x}$ 的邻域. 如果在定义 6.7 中只要求 (6.15), 也就是在 (6.16) 中 $\alpha = 0$, 那么可得到对应于 $C([a,b];X)$ 中范数拓扑下 $\bar{x}$ 邻域的经典强局部极小点. 如果 (6.16) 式换为更强的限制要求

$$\|\dot{x}(t) - \dot{\bar{x}}(t)\| < \varepsilon, \quad \text{a.e. } t \in [a,b],$$

那么在定义 6.7 的框架下得到经典的弱局部极小点. 这对应于考虑 $W^{1,\infty}([a,b];X)$ 中拓扑下的 $\bar{x}$ 邻域. 因此, 对任何 $p \in [1,\infty)$, 所引入的中间局部极小点 (i.l.m.) 的概念介于经典强 $(\alpha = 0)$ 和弱 $(p = \infty)$ 局部极小点概念之间. 很明显, 所有中间局

部极小点 (i.l.m.) 的必要条件对于强 (并因此对整体) 极小点自动成立. 下面在变分问题中考虑一些例子, 以此来阐明弱、中间和强局部极小点之间的关系.

第一个例子是标准的, 它表明即使在具有端点约束的最简单的经典变分法问题中, 弱和强极小点也可以是不同的.

例 6.8(弱但不是强的极小点) 存在经典的变分法问题, 其弱局部极小点不是强局部极小点.

证明 在满足端点约束 $x(0)=x(\pi)=0$ 的绝对连续函数 $x\colon[0,\pi]\to\mathbb{R}$ 上, 考虑变分问题:

$$\min \quad J[x]:=\int_0^{\pi} x^2(t)[1-\dot{x}^2(t)]\mathrm{d}t.$$

首先证明 $\bar{x}(\cdot)\equiv 0$ 是一个弱局部极小点. 事实上, 任取 $\varepsilon\in(0,1)$ 和任何可行轨线 $x\neq\bar{x}$ 满足

$$|x(t)-\bar{x}(t)|\leqslant\varepsilon,\ t\in[0,\pi], \quad 和 \quad |\dot{x}(t)-\dot{\bar{x}}(t)|\leqslant\varepsilon, \quad \text{a.e. } t\in[0,\pi],$$

有 $0<1-\varepsilon^2\leqslant 1-\dot{x}^2(t)$ 对几乎处处 $t\in[0,\pi]$ 成立. 因此, $x^2(t)[1-\dot{x}^2(t)]>0$ 对几乎处处 $t\in[0,\pi]$ 成立且 $J[x]>0=J[\bar{x}]$, 即 $\bar{x}$ 是一个弱局部极小点. 另一方面, $\bar{x}$ 不是强局部极小点, 这可由下面的证明得到. 对任何 $k\in\mathbb{N}$, 取可行轨线 $x_k(t):=(1/\sqrt{k})\sin(kt)$, 并注意到

$$J[x_k]=\frac{\pi}{2}\Big(\frac{1}{k}-\frac{1}{4}\Big)<0 \quad 对\ k\geqslant 5\ 成立,$$

然而, $|x_k(t)-\bar{x}(t)|\leqslant 1/\sqrt{k}$ 对所有 $t\in[0,\pi]$ 和 $k\in\mathbb{N}$ 成立. 因此, 给定任何 $\varepsilon>0$, 在 $C([0,\pi];\mathbb{R})$ 中总可找到一个属于 $\bar{x}$ 的 ε- 邻域的可行轨线 x_k, 使得 $J[x_k]<J[\bar{x}]$. △

接下来, 考虑一个不太标准的情形, 在定义 6.7 的意义下, 对任何秩为 $p\in[1,\infty)$ 的弱局部极小点, 它可能不是中间局部极小点. 该例也是一维框架下的经典变分法中的问题.

例 6.9(弱但不是中间的极小点) 存在一维变分法问题, 对任何秩 $p\geqslant 1$, 其弱极小点不是中间极小点.

证明 在满足端点约束 $x(0)=x(1)=0$ 的绝对连续函数 $x\colon[0,1]\to\mathbb{R}$ 上, 考虑变分问题:

$$\min \quad J[x]:=\int_0^1\big[\dot{x}^3(t)+3\dot{x}^2(t)\big]\mathrm{d}t.$$

为证 $\bar{x}(\cdot)\equiv 0$ 是一个弱局部极小点, 注意到, 只要 $\dot{x}(t)\geqslant -3$, 被积函数是非负的. 因此, 对任何满足

$$0<|\dot{x}(t)-\dot{\bar{x}}(t)|\leqslant\varepsilon<3, \quad \text{a.e. } t\in[0,1]$$

的可行轨线 x, 有 $J[x]>0$. 给定任何 $p\geqslant 1$, 现在证明 $\bar{x}$ 不是一个秩为 p 的中间局部极小点. 为此, 对任何自然数 $k\geqslant 3^{4p}$, 考虑一列可行轨线

$$x_k(t):=\begin{cases}-k^{\frac{1}{2p}}t, & 若\ 0\leqslant t\leqslant\dfrac{1}{k},\\ \dfrac{-k^{\frac{1}{2p}}(1-t)}{k-1}, & 若\ \dfrac{1}{k}<t\leqslant 1.\end{cases}$$

可验证

$$J[x_k]=-\frac{k^{\frac{1}{p}}}{(k-1)^2}\Big[(k^{\frac{1}{2p}}-3)(k-2)-3\Big]<0,$$

$$\int_0^1|\dot{x}_k(t)-\dot{\bar{x}}(t)|^p\mathrm{d}t=\frac{1}{\sqrt{k^p}}\Big(1+\frac{1}{(k-1)^{p-1}}\Big)^p\leqslant\Big(\frac{2}{\sqrt{k}}\Big)^p.$$

因此, 对任何 $\varepsilon>0$ 和任何 $p\geqslant 1$, 有

$$\int_0^1|\dot{x}_k(t)-\dot{\bar{x}}(t)|^p\mathrm{d}t\leqslant\varepsilon^p,\quad J[x_k]<0\quad\Big(\forall k\geqslant\max\{\varepsilon^{-2p},3^{4p}\}\Big),$$

这证明了 $\bar{x}$ 不可能是秩为 p 的中间极小点.

考虑上述问题的简化形式

$$\begin{aligned}&\min\quad J[x]:=\int_0^1\dot{x}^3(t)\mathrm{d}t\\&\text{s.t.}\quad x(0)=0,\quad x(1)=1.\end{aligned}$$

则轨线 $\bar{x}(t)=t$ 是一个弱局部极小点, 但不是一个秩为 $p\geqslant 2$ (但不是 $p\geqslant 1$) 的中间局部极小点. 为此, 取函数 $x_k(t)=\bar{x}(t)+y_k(t)$ 满足 $y_k(0)=y_k(1)=0$ 和

$$\dot{y}_k(t)=\begin{cases}-\sqrt{k}, & 若\ 0\leqslant t\leqslant\dfrac{1}{k},\\ \sqrt{k}(k-1)^{-1}, & 若\ \dfrac{1}{k}<t\leqslant 1\end{cases}$$

并且对每个 $p\in[2,\infty)$, 直接验证

$$J[x_k]=-\sqrt{k}+O(1)\to-\infty,\quad 然而\quad\int_0^1|\dot{x}_k(t)-\dot{\bar{x}}(t)|^p\mathrm{d}t\to 0\quad(k\to\infty),$$

讨论完毕. △

上述例子考虑没有微分包含/动态约束的变分法问题. 下一个例子涉及自治、凸值和 Lipschitz 微分包含并且表明强和中间局部极小点的概念在此情形可能不同.

例 6.10(对有界、凸值和 Lipschitz 微分包含, 中间但不是强的局部极小点) 在具有紧凸值的自治、一致有界和 Lipschitz 微分包含的轨道上, 存在一个极小化线

性费用函数的最优控制问题, 对任何秩 $p \in [1,\infty)$, 其上的中间局部极小点不是强局部极小点.

证明　设 $x = (x_1, x_2, x_3, x_4) \in \mathbb{R}^4$, 并且

$$\psi(x_1, x_2) := \begin{cases} x_2^2 \cos\left(\dfrac{\pi x_1}{x_2}\right), & 若\ x_2 \neq 0, \\ 0, & 若\ x_2 = 0. \end{cases}$$

容易验证 ψ 在 $\mathbb{R}^4$ 上是连续可微的. 在具有端点约束

$$x_1(0) = x_4(0) = x_4(1) = 0, \quad x_1(1) = 1$$

的微分包含

$$\begin{pmatrix} \dot{x}_1 \\ \dot{x}_2 \\ \dot{x}_3 \\ \dot{x}_4 \end{pmatrix} \in \left\{ \left[\begin{array}{c} 1 \\ 0 \\ v \\ |\psi(x_1, x_2) - x_2 x_3| \end{array} \right] \middle| v \in [-4, 4] \right\}, \quad \text{a.e. } t \in [0, 1]$$

的绝对连续轨道上考虑下列问题:

$$\min \quad J[x] := -x_2(1).$$

取一个可行轨线 $\bar{x}(t) = (t, 0, 0, 0)$, 并且首先证明它不是强局部极小点. 事实上, 对任何 $\varepsilon \in (0, 2\sqrt{2})$, 函数

$$x(t) = \left(t, \frac{\varepsilon}{\sqrt{2}}, \frac{\varepsilon}{\sqrt{2}} \cos\left(\frac{\sqrt{2}\pi t}{\varepsilon}\right), 0\right)$$

在空间 $C([0,1];\mathbb{R}^4)$ 的 $\bar{x}$ 的 ε- 邻域上是一个可行轨线, 并且 $J[x] = -\varepsilon/\sqrt{2} < 0 = J[\bar{x}]$.

其次证明对秩 $p = 1$, $\bar{x}$ 是一个中间局部极小点. 因此, 对任何秩 $p \in [1,\infty)$, $\bar{x}$ 也是一个中间局部极小点. 任取 $\varepsilon \in (0, 1/2)$, 并相反地假设存在一个可行轨线 $x(\cdot)$ 满足定义 6.7 中的关系式 (6.15) 和 (6.16), 使得 $J[x] < J[\bar{x}]$. 那么在 $[0,1]$ 上对某 $c \in (0, 1/2)$ 有

$$x_1(t) = t, \quad x_2(t) \equiv c, \quad |\ \psi(t, c) - c x_3(t)\ | \equiv 0.$$

由此有

$$x_3(t) = \frac{\psi(t, c)}{c} = c \cos\left(\frac{\pi t}{c}\right), \quad 因此\ \dot{x}_3 = -\pi \sin\left(\frac{\pi t}{c}\right).$$

由于 $c \in (0, 1/2)$, 故有

$$\int_0^1 \|\dot{x}(t) - \dot{\bar{x}}(t)\| \mathrm{d}t = \pi \int_0^1 \left|\sin\left(\frac{\pi t}{c}\right)\right| \mathrm{d}t = \pi c \int_0^{c^{-1}} |\sin(\pi s)| \mathrm{d}s$$
$$\geqslant \pi c \int_0^{[c^{-1}]} |\sin(\pi s)| \mathrm{d}s = 2c\left[\frac{1}{c}\right] \geqslant \frac{2}{3},$$

其中 $[a]$ 表示小于或等于 $a \in \mathbb{R}$ 的最大整数. 这显然与 $\varepsilon < 1/2$ 的选取相矛盾. 这就证明了 $\bar{x}$ 是秩为 $p = 1$ 的中间局部极小点. △

接下来, 与原问题 (P) 一道, 考虑其相应的松弛问题, 粗略地讲, 它是由 (P) 通过关于速率变量的凸化得到. 对 (6.13) 式中的被积函数 $\vartheta(x, v, t)$, 考虑它在 (6.1) 式中集合 $F(x, t)$ 上的限制

$$\vartheta_F(x, v, t) := \vartheta(x, v, t) + \delta(v; F(x, t)),$$

并记 $\widehat{\vartheta}_F(x, v, t)$ 为 $\vartheta_F(x, \cdot, t)$ 的二次共轭 (双极化) 函数, 即

$$\widehat{\vartheta}_F(x, v, t) = (\vartheta_F)_v^{**}(x, v, t), \quad \forall (x, v, t) \in X \times X \times [a, b].$$

众所周知, $\widehat{\vartheta}_F(x, v, t)$ 是关于 v 的最大正常、凸下半连续函数, 它由 ϑ_F 界定. 此外, $\vartheta_F = \widehat{\vartheta}_F$ 当且仅当 ϑ_F 关于 v 是正常、凸和下半连续的.

给定原变分问题 (P), 与 $\vartheta_F(x(t), \dot{x}(t), t)$ 一道, 对于 $[a, b]$ 上几乎处处可微的轨线 $x \colon [a, b] \to X$, 它在 $[a, b]$ 上 Bochner 可积并且满足 Newton-Leibniz 公式及端点约束 (6.14), 定义松弛问题 (R) (或 (P) 的松弛) 如下:

$$\min \quad \widehat{J}[x] := \varphi(x(a), x(b)) + \int_a^b \widehat{\vartheta}_F(x(t), \dot{x}(t), t) \mathrm{d}t. \tag{6.17}$$

注意到, 与 (6.13) 不同的是, (6.17) 中的被积函数是增广实值的. 更进一步, 在自然的要求 $\widehat{J}[x] < \infty$ 下, $x(\cdot)$ 是凸化微分包含

$$\dot{x}(t) \in \mathrm{clco} F(x(t), t), \quad \text{a.e. } t \in [a, b] \tag{6.18}$$

的一个解 (在定义 6.1 的意义下). 因此, 松弛问题 (R) 的动态约束是由凸化微分包含 (6.18) 显式给出的. (6.18) 的任何一个轨道称为 (6.1) 的一个松弛轨道.

在微分包含的松弛和原轨道之间有着深刻的关系, 它反映了由微分算子所定义的连续时间 (无原子测度) 的动态系统中固有的隐含凸性. 在本章随后的材料中会看到这一现象的各种实现. 尤其, Banach 空间中任何紧值和关于 x 是 Lipschitz 的微分包含的松弛轨道可以 (在空间 $C([a, b]; X)$ 中) 由从相同初始状态 $x(a) = x_0$ 开始的最初轨道一致逼近; 例如参见 Tolstonov [1258] 中的定理 2.2.1 和其中的参考

文献. 这里需要的这个逼近/稠密性质的形式不仅涉及微分包含而且也涉及极小化泛函. 当 Banach 空间可分时, 下面的结果成立, 其证明由 De Blasi, Pianigiani 和 Tolstonogov[308] 给出. 这种类型的结果可追溯到 Bogolyubov[121] 和 Young[1350] 在变分法中的经典定理.

定理 6.11(松弛轨道的逼近性质) 设 $x(\cdot)$ 为微分包含 (6.1) 的一个松弛轨道, X 是可分的并且 $F\colon X\times[a,b]\rightrightarrows X$ 是紧值并以可和函数为界是一致有界的、关于 x 是局部 Lipschitz 且关于 t 是可测的. 假设 (6.13) 中的被积函数 ϑ 关于 (x,v) 连续, 关于 t 可测并且在 $x(\cdot)$ 附近以可和函数为界是一致有界的. 那么存在 (6.1) 的一列原始轨道 $x_k(\cdot)$ 满足关系

$$x_k(a)=x(a),\quad x_k(\cdot)\to x(\cdot)\quad (\text{在 } C([a,b];X) \text{ 中}),$$

$$\liminf_{k\to\infty}\int_a^b\vartheta(x_k(t),\dot{x}_k(t),t)\mathrm{d}t\leqslant\int_a^b\widehat{\vartheta}_F(x_k(t),\dot{x}_k(t),t)\mathrm{d}t.$$

注意到, 定理 6.11 并不要求近似轨道 $x_k(\cdot)$ 满足端点约束 (6.14). 事实上, 有例子表明这可能不成立. 如果成立的话, 那么问题 (P) 具有松弛稳定性:

$$\inf(P)=\inf(R), \tag{6.19}$$

其中费用泛函 (6.13) 和 (6.17) 的下确界是在 (P) 和 (R) 的所有可行轨道上分别取的.

松弛稳定性的一个明显的充分条件是集合 $F(x,t)$ 和被积函数 ϑ 关于 v 的凸性. 然而, 由于连续时间微分包含的隐含凸性, 松弛稳定性要比标准凸性深刻得多. 尤其, 定理 6.11 确保没有端点约束 $x(b)$ 的非凸问题 (P) 的松弛稳定性. 对于非凸问题的松弛稳定性还有其他的有效条件, 具体讨论举例可见 Ioffe 和 Tikhomirov [617], Mordukhovich [888, 915] 和 Tolstonogov [1258]. 值得一提的结果/概念包括, 经典的 Bogolyubov 定理, 它确保了在具有端点约束 (6.14) 但无动态约束 (6.1) 的极小化 (6.13) 的变分问题的松弛稳定性; 由关于无原子向量测度值域凸性的 Lyapunov 定理 (这在很大程度上表明了隐含的凸性) 得到的关于状态变量是线性的控制系统的松弛稳定性; 由 Clarke[246, 255] 对具有不等式端点约束的微分包含提出的平静性条件; 由 Warga[1315, 1321] 提出的涉及参数控制系统 (6.2) 的正规性条件等.

这里研究的本质部分涉及关于松弛稳定的局部极小值. 对应定义 6.7, 有如下定义.

定义 6.12(松弛中间局部极小点) 称轨线 $\bar{x}$ 是原问题 (P) 的一个秩为 $p\in[1,\infty)$ 的松弛中间局部极小点 (r.i.l.m.), 如果 $\bar{x}$ 是 (P) 的可行解且是具有相同费用函数 $J[\bar{x}]=\widehat{J}[\bar{x}]$ 的松弛问题 (R) 的具有该秩的中间局部极小点.

类似地, 可定义松弛弱和松弛强局部极小点, 它们之间具有上述讨论的相同关系. 当然, 对具有凸集 $F(x,t)$ 和关于速率为凸的被积函数的问题 (P), 其局部极小点之间相应的松弛和通常 (非松弛) 的概念是没有区别的. 显然, 问题 (P) 的任何松弛中间 (弱、强) 极小点为原始问题提供了一个相应的非松弛局部极小点. 反之, 则需要一种局部松弛稳定性. 注意到, 松弛中间局部极小点的任何必要条件对松弛强局部极小点也成立. 因此, 在松弛稳定性 (6.19) 下, 对 (P) 的最优解 (整体或绝对极小点) 也成立.

这里的主要目标是对所考虑的 Bolza 问题 (P) 导出松弛中间局部极小点的一般必要最优性条件; 其中的一些结果也会在以后无任何松弛下得到. 为此, 利用离散逼近的方法, 该方法把连续时间系统的变分/最优控制问题及其相应的有限差分版本联系在一起. 该方法的第一步是对问题 (P) 中给定的松弛中间局部极小点 $\bar{x}(\cdot)$ 构造适定的离散逼近, 使得离散时间问题的最优解在 $W^{1,2}([a,b];X)$ 空间中强收敛于 $\bar{x}(\cdot)$. 这将在下面的小节中完成.

6.1.3 Bolza 问题的适定离散逼近

对于 6.1.1 小节考虑的微分包含及其对应的有限差分变体, 那里, 在一般的 Banach 空间中, 建立了微分包含的每一个轨道可由有限差分包含在自然假设下的扩张轨道强逼近. 该结果虽然没有直接关联到涉及微分包含的最优化问题, 但是这里将在最优化的框架下利用它. 该小节的主要目标如下.

给定微分包含 (6.1) 的一个轨道 $\bar{x}(\cdot)$, 该轨道为上面定义的最优化问题 (P) 提供了一个松弛中间局部极小点. 下面对有限差分包含 (6.3) 构造一族适定最优化问题 (P_N), 使得 (P_N) 的 (扩张) 最优解 $\bar{x}_N(\cdot)$ 在 $W^{1,2}([a,b];X)$ 中的范数拓扑下强收敛于 $\bar{x}(\cdot)$.

在 6.1.1 小节的常驻假设 (H1) 和 (H2) 下, 注意到有界性假设 (6.4) 蕴涵定义 6.12 的松弛中间局部极小点的概念并不依赖秩 $p\in[1,\infty)$. 这意味着, 如果 $\bar{x}(\cdot)$ 是对某秩 $p\in[1,\infty)$ 的一个松弛中间局部极小点, 那么它也是对任何其他秩 $p\geqslant 1$ 的一个松弛中间局部极小点. 为简单起见, 下面取 $p=2$ 和 $\alpha=1$.

为此, 除了那些附加在 F 上的假设外, 需要对问题 (P) 中的其他数据 ϑ, φ 和 Ω 作适当假设. 因为涉及 Bochner 积分, 所以总是将映射 $f\colon[a,b]\to X$ 的可测性等同于强可测性, 其中 f 是强可测的, 如果它在 $[a,b]$ 的可测子集上能被一列取值于 X 的阶梯函数几乎处处逼近. 给定 $\bar{x}(\cdot)$ 的一个邻域 U 和 (H1) 中的常数 m_F, 进一步假设:

(H3) $\vartheta(\cdot,\cdot,t)$ 在 $U\times(m_F\mathbb{B})$ 上连续, 关于 $t\in[a,b]$ 是一致的; 同时 $\vartheta(x,v,\cdot)$ 在 $[a,b]$ 上可测并且其范数被可和函数界定, 关于 $(x,v)\in U\times(m_F\mathbb{B})$ 是一致的.

(H4) φ 在 $U\times U$ 上连续; $\Omega\subset X\times X$ 在 $(\bar{x}(a),\bar{x}(b))$ 附近是局部闭的并且

对某 $\varepsilon > 0$, 集合 $\mathrm{proj}_1\Omega \cap (\bar{x}(a) + \varepsilon\mathbb{B})$ 是紧的, 其中 $\mathrm{proj}_1\Omega$ 表示 Ω 在乘积空间 $X \times X$ 中的第一个空间 X 上的投影.

对称地, 可以假设 $\Omega \subset X \times X$ 在第二个空间上的投影是局部紧的; 为了明确起见, (H4) 中选取在第一个空间上的投影.

现在考虑松弛中间局部极小点 $\bar{x}(\cdot)$, 对这个可行轨线, 应用关于离散轨道强逼近的定理 6.4 . 于是找到一列逼近 $\bar{x}(\cdot)$ 的扩张离散轨道 $\widehat{x}_N(\cdot)$ 并且计算 (6.11) 中的数值 η_N. 由于从关于 $p = 2$ 和 $\alpha = 1$ 的中间局部极小点 $\bar{x}(\cdot)$ 的关系式 (6.15) 和 (6.16) 中有 $\varepsilon > 0$, 对所有 $t \in [a,b]$, 总假设 $\bar{x}(t)+(\varepsilon/2)\mathbb{B} \in U$. 构造一列离散逼近问题 (P_N), $N \in \mathbb{N}$ 如下: 在差分包含 (6.3) 的离散轨道 $x_N = x_N(\cdot) = (x_N(t_0), \cdots, x_N(t_N))$ 上, 极小化离散时间 Bolza 泛函

$$\begin{aligned} J_N[x_N] := &\varphi(x_N(t_0), x_N(t_N)) + \|x_N(t_0) - \bar{x}(a)\|^2 \\ &+ \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t\right) \mathrm{d}t \\ &+ \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 \mathrm{d}t, \end{aligned} \tag{6.20}$$

其中约束为

$$(x_N(t_0), x_N(t_N)) \in \Omega + \eta_N\mathbb{B}, \tag{6.21}$$

$$\|x_N(t_j) - \bar{x}(t_j)\| \leqslant \frac{\varepsilon}{2}, \quad j = 1, \cdots, N, \tag{6.22}$$

$$\sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 \mathrm{d}t \leqslant \frac{\varepsilon}{2}. \tag{6.23}$$

如 6.1.1 小节一样, 考虑 $x_N(\cdot)$ 在整个区间 $[a,b]$ 上的具有分段常数导数的分段线性扩张 (以后不再赘述). 于是有

$$\begin{cases} x_N(t) = x_N(a) + \displaystyle\int_a^t \dot{x}_N(s)\mathrm{d}s, \quad \forall\, t \in [a,b], \\ \dot{x}_N(t) = \dot{x}_N(t_j) \in F(x_N(t_j), t_j), \quad t \in [t_j, t_{j+1}), \quad j = 0, \cdots, N-1. \end{cases} \tag{6.24}$$

下一个定理建立了 (P) 的局部极小点 $\bar{x}(\cdot)$ 能被 (P_N) 的最优解在 $W^{1,2}([a,b];X)$ 中强逼近, 这意味着导数的几乎处处逐点收敛, 而这在以下的讨论中是至关重要的. 为了证明这样的一个逼近, 需要所论空间 X 的 Asplund 结构和 Radon-Nikodým性质 (RNP), 这样就能确保可以利用关于 $L^1([a,b];X)$ 中集合弱紧性的经典 Dunford 定理. 值得注意的是: 每一个自反空间都是 Asplund 的并且具有 RNP. 更进一步, 如果 X^* 是 Asplund 的, 那么它的二次对偶空间 X^{**} 具有 RNP (于是 $X \subset X^{**}$ 也

具有 RNP). 因此, Banach 空间 X, 其中的 X 和 X^* 都是 Asplund 的, 满足下一个定理所需的性质. 正如在 3.2.5 小节开始讨论的那样, 存在非自反 (甚至是可分的) 空间属于这类空间.

定理 6.13(离散最优解的强逼近) 在假设 (H1)~(H4) 下, 设 $\bar{x}(\cdot)$ 为 Bolza 问题 (P) 的松弛中间局部极小点并且设 (P_N), $N \in \mathbb{N}$ 为上面所构造的一列离散逼近问题. 那么下列成立:

(i) 每个 (P_N) 都具有一个最优解.

(ii) 如果进一步 X 是 Asplund 的并且具有 RNP, 那么 (P_N) 的任何一列最优解 $\{\bar{x}_N(\cdot)\}$ 在 $W^{1,2}([a,b];X)$ 中强收敛于 $\bar{x}(\cdot)$.

证明 为证 (i), 注意到对充分大的 N, 问题 (P_N) 的可行轨道集是非空的, 这是因为定理 6.4 中的扩张函数 $\widehat{x}_N(\cdot)$, 根据其构造, 满足 (6.3) 和约束 (6.21)~(6.23). 对于 (6.21) 和 (6.22) 的情形, 由 (6.11) 即可得. 在 (6.23) 的情形中, 对充分大的 N, 由 (6.4) 和 (6.12) 可得

$$\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\left\|\frac{\widehat{x}_N(t_{j+1})-\widehat{x}_N(t_j)}{h_N}-\dot{\bar{x}}(t)\right\|^2\mathrm{d}t=\int_a^b\|\dot{\widehat{x}}_N(t)-\dot{\bar{x}}(t)\|^2\mathrm{d}t$$

$$\leqslant 2m_F\alpha_N\leqslant\frac{\varepsilon}{2},$$

其中 α_N 在 (6.12) 式中定义. 由 (H1), (H3) 和 (H4) 中对紧性和连续性所作的假设, (P_N) 的最优解的存在性由经典的 Weierstrass 定理可得.

剩下证明收敛性的论断 (ii). 首先沿着 $N \in \mathbb{N}$ 的某序列验证

$$J_N[\widehat{x}_N]\to J[\bar{x}]\quad(N\to\infty).\tag{6.25}$$

考虑到 (6.20) 中关于 $J_N[\widehat{x}_N]$ 的表达式, 由定理 6.4 可推断出其中的第二项趋于零, 由 (6.4) 和 (6.2) 第四项趋于零并且由 (H4) 中关于 φ 的连续性的假设, 第一项趋于 $\varphi(\bar{x}(a),\bar{x}(b))$. 因此, 只需证明当 $N\to\infty$ 时,

$$\sigma_N:=\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\vartheta\left(\widehat{x}_N(t_j),\frac{\widehat{x}_N(t_{j+1})-\widehat{x}_N(t_j)}{h_N},t\right)\mathrm{d}t\to\int_a^b\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t)\mathrm{d}t.$$

使用记号“$\sim$”表示表达式当 $N\to\infty$ 时是等价的. 定理 6.4 确保沿着 $N\to\infty$ 的某一子列, $\dot{\widehat{x}}_N(t)\to\dot{\bar{x}}(t)$ 是几乎处处的并且在 (H3) 的假设下, 关于 Bochner 积分的 Lebesgue 控制收敛定理是成立的. 于是有

$$\sigma_N=\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\vartheta(\widehat{x}_N(t_j),\dot{\widehat{x}}_N(t),t)\mathrm{d}t\sim\int_a^b\vartheta(\widehat{x}_N(t),\dot{\widehat{x}}_N(t),t)\mathrm{d}t$$

$$\sim\int_a^b\vartheta(\bar{x}(t),\dot{\widehat{x}}_N(t),t)\mathrm{d}t\sim\int_a^b\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t)\mathrm{d}t.$$

注意到, 在任意 Banach 空间 X 中, 对原问题 (P) 的任何中间 (不是松弛) 局部极小点 $\bar{x}(\cdot)$, 已证明了 (6.25). 接下来证明如果 $\bar{x}(\cdot)$ 是原问题的一个松弛中间局部极小点, 其中的状态空间 X 是 Asplund 的并且满足 RNP, 那么对于 (P_N) 的每一列最优解 $\bar{x}_N(\cdot)$, (6.25) 式蕴涵

$$\lim_{N\to\infty}\left[\beta_N := \|\bar{x}_N(a)-\bar{x}(a)\|^2+\int_a^b\|\dot{\bar{x}}_N(t)-\dot{\bar{x}}(t)\|^2\mathrm{d}t\right]=0. \tag{6.26}$$

假设 (6.26) 式不成立, 那么在 (6.26) 式中选取序列 $\{\beta_N\}$ 的一个极限点 $\beta>0$, 为方便起见, 假设 $\beta_N\to\beta$ 对所有 $N\to\infty$. 接下来, 在空间 $L^1([a,b];X)$ 中, 对序列 $\{\dot{\bar{x}}_N(\cdot)\}$, $N\in\mathbb{N}$ 利用关于相对弱紧的 Dunford 定理 (参见 Diestel 和 Uhl [334, 定理 IV.1]). 由 (6.24) 和 (H1), 该序列满足 Dunford 定理的条件. 更进一步, 空间 X 和 X^* 都具有 RNP, 这是因为如上所述 X^* 的 RNP 等价于 X 的 Asplund 结构. 因此, 不失一般性, 假设存在 $v\in L^1([a,b];X)$, 使得

$$\dot{\bar{x}}_N(\cdot)\to v(\cdot)\quad 当\ N\to\infty\ 时在\ L^1([a,b];X)\ 中弱收敛.$$

因为 Bochner 积分是从 $L^1([a,b];X)$ 到 X 的线性连续算子, 所以它仍然连续, 若空间 $L^1([a,b];X)$ 和 X 装备弱拓扑. 由 (6.21) 和 (H4) 中关于 Ω 的假设, 集合 $\{\bar{x}_N(a)\mid N\in\mathbb{N}\}$ 在 X 中是相对紧的. 由 (6.24) 和微分包含解的紧性 (这在 (H1) 和 (H2) 的假设下可得到; 参见例如 Tolstonogov [1258, 定理 3.4.2]), 可断言序列 $\{\bar{x}_N(\cdot)\}$ 包含一个子列, 该子列在空间 $C([a,b];X)$ 中依范数拓扑收敛于 $\widetilde{x}(\cdot)$. 现在在关于 $\bar{x}_N(\cdot)$ 的 Newton-Leibniz 公式 (6.24) 中取极限并且考虑到上述收敛性, 于是有

$$\widetilde{x}(t)=\widetilde{x}(a)+\int_a^b v(s)\mathrm{d}s,\quad \forall\, t\in[a,b],$$

这意味着 $\widetilde{x}(\cdot)$ 在 $[a,b]$ 上的绝对连续性和几乎处处可微性并且 $v(t)=\dot{\widetilde{x}}(t)$, a.e. $t\in[a,b]$. 注意到 $\widetilde{x}(\cdot)$ 是凸化微分包含 (6.18) 的一个解. 事实上, 因为 $\{\dot{\bar{x}}_N(\cdot)\}$ 的一个子列在 $L^1([a,b];X)$ 中弱收敛于 $\dot{\widetilde{x}}(\cdot)$. 因此, $\dot{\bar{x}}_N(\cdot)$ 的某凸组合沿 $L^1([a,b];X)$ 的范数拓扑收敛到 $\dot{\widetilde{x}}(\cdot)$, 进而对几乎处处 $t\in[a,b]$ 是逐点收敛的. 在关于 $\bar{x}_N(\cdot)$ 的微分包含 (6.24) 中取极限, 可得 $\widetilde{x}(\cdot)$ 满足 (6.18). 在 (6.21) 和 (6.22) 中取极限, 也可得到 $\widetilde{x}(\cdot)$ 满足端点约束 (6.14) 和

$$\|\widetilde{x}(t)-\bar{x}(t)\|\leqslant\varepsilon/2,\quad \forall\, t\in[a,b].$$

更进一步, 积分泛函

$$I[v]:=\int_a^b\|v(t)-\dot{\bar{x}}(t)\|^2\mathrm{d}t$$

在 $L^2([a,b];X)$ 的弱拓扑下是下半连续的, 这是由于被积函数关于 v 的凸性. 因为由 (6.4) 中有界性的假设, $L^1([a,b];X)$ 中的弱收敛 $\dot{\bar{x}}_N(\cdot) \to \dot{\tilde{x}}(\cdot)$ 蕴涵 $L^2([a,b];X)$ 中的弱收敛, 并且因为

$$\int_a^b \|\dot{\bar{x}}_N(t) - \dot{\bar{x}}(t)\|^2 \mathrm{d}t = \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 \mathrm{d}t,$$

上述下半连续性和关系式 (6.23) 蕴涵

$$\int_a^b \|\dot{\tilde{x}}(t) - \dot{\bar{x}}(t)\|^2 \mathrm{d}t \leqslant \liminf_{N\to\infty} \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\|^2 \mathrm{d}t \leqslant \frac{\varepsilon}{2}.$$

因此, 在空间 $W^{1,2}([a,b];X)$ 中, 轨线 $\tilde{x}(\cdot)$ 属于 $\bar{x}(\cdot)$ 的 ε- 邻域.

最后证明对费用泛函 (6.17), 轨线 $\tilde{x}(\cdot)$ 给出比 $\bar{x}(\cdot)$ 更小的值. 因为每个 $\widehat{x}_N(\cdot)$ 对 (P_N) 是可行的, 所以总有

$$J_N[\bar{x}_N] \leqslant J_N[\widehat{x}_N], \quad \text{对所有充分大的 } N \in \mathbb{N} \text{成立}.$$

现在当 $N \to \infty$ 时取极限并且考虑到上述的讨论以及 (6.17) 中凸化被积函数 $\widehat{\vartheta}_F$ 的构造, 于是由 (6.25) 有

$$\varphi(\tilde{x}(a), \tilde{x}(b)) + \int_a^b \widehat{\vartheta}_F(\tilde{x}(t), \dot{\tilde{x}}(t), t)\mathrm{d}t + \beta \leqslant J[\bar{x}].$$

由于 $\beta > 0$, 这样就得到 $\widehat{J}[\tilde{x}] < J[\bar{x}] = \widehat{J}[\bar{x}]$. 这是不可能的, 因为 $\bar{x}(\cdot)$ 是 (P) 的一个松弛中间局部极小点. 于是 (6.26) 成立, 这显然蕴涵着在空间 $W^{1,2}([a,b];X)$ 的范数拓扑下的收敛 $\bar{x}_N(\cdot) \to \bar{x}(\cdot)$. 定理证毕. △

由定理 6.13 证明所展述的论证, 可以建立离散逼近值收敛的有效条件; 这种收敛的意思是离散逼近问题中费用泛函的最优/最小值收敛于原问题中费用泛函的最优/最小值. 此外, 利用定理 6.11 中松弛轨道的逼近性质实际上可得到由原问题的一个内在性质表述的值收敛的一个必要充分条件.

注意到离散问题 (P_N) 中的费用泛函 (6.20) 和约束 (6.22) 及 (6.23) 显式地包含 (P) 的给定极小点 $\bar{x}(\cdot)$. 下面, 考虑到离散逼近的值收敛, 在离散逼近的构造中将不涉及任何局部极小值并且/或者甚至不假设原问题最优解的存在性. 为此, 考虑一列新的离散逼近问题 $(\widetilde{P}_N)$, 其构造如下:

$$\begin{aligned} \min \widetilde{J}_N[x_N] := & \varphi(x_N(t_0), x_N(t_N)) \\ & + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(x_N(t_j), \frac{x_N(t_{j+1}) - x_N(t_j)}{h_N}, t\right) \mathrm{d}t, \end{aligned}$$

其中约束包括离散包含 (6.3) 和扰动的端点约束 (6.21), 序列 η_N 待以后确定. 注意到, 问题 $(\widetilde{P}_N)$ 构造性地建立在初始连续时间的初始数据之上. 在下一个定理中, 记号 $\widetilde{J}_N^0 := \inf(\widetilde{P}_N)$, $\inf(P)$ 和 $\inf(R)$ 分别代表问题 $(\widetilde{P}_N)$, (P) 和 (R) 中费用泛函的最优值. 注意到, 离散时间问题 $(\widetilde{P}_N)$ 的最优解总是存在的, 这是由于定理 6.13 中所作的假设 (H1)~(H4), 其中需要端点约束的适当扰动 η_N; 参见其证明.

定理 6.14(离散逼近的值收敛) 设 $U \subset X$ 为 Banach 空间的一个开子集, 使得对 (P) 的可行解的一个极小序列, 当 $t \in [a,b]$ 和 $k \in \mathbb{N}$ 时, 有 $x_k(t) \in U$. 假设在集合 U 上条件 (H1)~(H4) 成立, 其中 (H4) 中 $\bar{x}(a) + \varepsilon\mathbb{B}$ 由 $\mathrm{cl}\,U$ 替换, 则下列论断成立:

(i) 在 (6.21) 中存在一列端点约束扰动 $\eta_N \downarrow 0$, 使得

$$\inf(R) \leqslant \liminf_{N\to\infty} \widetilde{J}_N^0 \leqslant \limsup_{N\to\infty} \widetilde{J}_N^0 \leqslant \inf P, \tag{6.27}$$

其中左端的不等式要求 X 是 Asplund 的并且具有 RNP. 因此, 如果 X 是 Asplund 的并且具有 RNP, 那么 (P) 的松弛稳定性 (6.19) 对于离散逼近的值收敛

$$\inf(\widetilde{P}_N) \to \inf(P) \quad (N \to \infty)$$

是充分的.

(ii) 反之, 如果 X 是自反和可分的, 那么对于端点约束的任何扰动 $\eta_N \downarrow 0$, (P) 的松弛稳定性也是离散逼近值收敛 $\inf(\widetilde{P}_N) \to \inf(P)$ 的一个必要条件.

证明 先证明在任何 Banach 空间中, (6.27) 右端的不等式成立. 选取定理中提到的 (P) 的可行轨道的极小序列 $x_k(\cdot)$, 对每个 $x_k(\cdot)$ 利用关于离散轨道强逼近的定理 6.4. 由对角化过程, 构建 (6.3) 的扩张离散轨道 $\widehat{x}_N(\cdot)$, 使得

$$\eta_N := \|(\widehat{x}_N(a), \widehat{x}_N(b)) - (x_{k_N}(a), x_{k_N}(b))\| \to 0 \quad (N \to \infty)$$

并且对于 (6.21) 中的这些约束扰动 η_N, 考虑离散逼近问题序列 $(\widetilde{P}_N)$. 类似定理 6.13 证明的第一部分, 证明每个 $(\widetilde{P}_N)$ 有一个最优解且利用反证法证明 (6.27) 右端的不等式. 为验证 (6.27) 左端的不等式, 假设 X 是 Asplund 的且具有 RNP, 并遵照定理 6.13 第二部分的证明. 在 (P) 的松弛稳定性之下, 这自动蕴涵值收敛 $\inf(\widetilde{P}_N) \to \inf(P)$.

为证定理中相反的论断 (ii), 首先注意到, 在所作的假设之下, 松弛问题 (R) 有一个最优解; 参见 Tolstonogov [1258, 定理 A.1.3]. 由定理 6.13 第二部分的讨论, 如果 X 是 Asplund 的且具有 RNP, 在可行解 $x(\cdot)$ 的集合上由导数 $\dot{x}(\cdot) \in L^1([a,b];X)$ 的弱收敛所诱导的拓扑之下, 实际上证明了松弛问题可行解的紧性和极小化泛函 (6.17) 的下半连续性. 现在假设 X 是自反和可分的并且利用定理 6.11, 由收敛

于 $\bar{x}(\cdot)$ 的初始轨道列 $x_k(\cdot)$ 逼近某个松弛最优轨道 $\bar{x}(\cdot)$. 接下来, 由定理 6.4, 在 $W^{1,2}([a,b];X)$ 中, 每个 $x_k(\cdot)$ 能被离散轨道 $\widehat{x}_{k_N}(\cdot)$ 强逼近. 利用对角化过程, 可得到一列逼近 $\bar{x}(\cdot)$ 的离散轨道 $\widehat{x}_N(\cdot)$ 并且令

$$\eta_N := \|(\widehat{x}_N(a), \widehat{x}_N(b)) - (\bar{x}(a), \bar{x}(b))\| \to 0 \quad (N \to \infty).$$

现在假设问题 (P) 关于松弛是不稳定的, 即 $\inf(R) < \inf(P)$, 并且对于一列具有端点约束 (6.21) 的某些扰动 η_N 的离散逼近问题 $(\widetilde{P}_N)$ 证明

$$\liminf_{N\to\infty} \widetilde{J}_N^0 < \inf(P).$$

事实上, 对于松弛最优轨道 $\bar{x}(\cdot)$, 有

$$\inf(R) = \varphi(\bar{x}(a), \bar{x}(b)) + \int_a^b \widehat{\vartheta}_F(\bar{x}(t), \dot{\bar{x}}(t), t)\mathrm{d}t < \inf(P),$$

构造 η_N 如上并且考虑具有这些扰动的端点约束问题 $(\widetilde{P}_N)$. 考虑到在定理 6.11 给出的由 $x_k(\cdot)$ 逼近的 $\bar{x}(\cdot)$, 在定理 6.4 中由离散轨道 $\widehat{x}_N(\cdot)$ 强逼近 $x_k(\cdot)$ 以及关系式

$$\begin{aligned}\widetilde{J}_N^0 &\leqslant \varphi(\widehat{x}_N(t_0), \widehat{x}_N(t_N)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta\left(\widehat{x}_N(t_j), \frac{\widehat{x}_N(t_{j+1}) - \widehat{x}_N(t_j)}{h_N}, t\right)\mathrm{d}t\\ &= \varphi(\widehat{x}_N(a), \widehat{x}_N(b)) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta(\widehat{x}_N(t_j), \dot{\widehat{x}}_N(t), t)\mathrm{d}t,\end{aligned}$$

在没有松弛稳定性之下有

$$\begin{aligned}\liminf_{N\to\infty} \widetilde{J}_N^0 &\leqslant \liminf_{N\to\infty}\left[\varphi(\widehat{x}_N(a), \widehat{x}_N(b)) + \int_a^b \vartheta(\widehat{x}_N(t), \dot{\widehat{x}}_N(t), t)\mathrm{d}t\right]\\ &\leqslant \varphi(\bar{x}(a), \bar{x}(b)) + \int_a^b \widehat{\vartheta}_F(\bar{x}(t), \dot{\bar{x}}(t), t)\mathrm{d}t < \inf(P).\end{aligned}$$

因此, 没有相应于端点约束上述扰动的离散逼近问题 $(\widetilde{P}_N)$ 的值收敛. 这就证明了 (ii). 证毕. △

因此, 在 (6.21) 中端点约束的适当扰动之下, (P) 的松弛稳定性作为连续时间动态最优问题的一个内在和自然的性质实际上是离散逼近值收敛的一个刻画. 由定理 6.14 的证明, 如果 (P) 有一个最优解 $\bar{x}(\cdot)$ 使得其导数 $\dot{\bar{x}}(\cdot)$ 在 $[a,b]$ 上 Riemann 可积, 那么相应的扰动 η_N 可通过连续性均模 (6.6) 表示为

$$\eta_N = \tau(\dot{\bar{x}}; h_N) \to 0 \quad (N \to \infty).$$

此外, 如果 $\dot{\bar{x}}(t)$ 在该区间上是有界变差的, 那么 $\eta_N = O(h_N)$; 参见 6.1.1 小节.

注 6.15(离散逼近的简化形式)　注意到, 如果 $\vartheta(x,v,\cdot)$ 在 $[a,b]$ 上是几乎处处连续的, 在最优解 $\bar{x}(\cdot)$ 的某个邻域内关于 (x,v) 是一致的, 那么在定理 6.13 中, 问题 (P_N) 的费用泛函 (6.20) 可替换为

$$\begin{aligned} J_N[x_N]:=&\varphi(x_N(t_0),x_N(t_N))+\|x_N(t_0)-\bar{x}(a)\|^2\\ &+h_N\sum_{j=0}^{N-1}\vartheta\left(x_N(t_j),\frac{x_N(t_{j+1})-x_N(t_j)}{h_N},t_j\right)\\ &+\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}}\left\|\frac{x_N(t_{j+1})-x_N(t_j)}{h_N}-\dot{\bar{x}}(t)\right\|^2\mathrm{d}t; \end{aligned} \tag{6.28}$$

并且定理 6.14 中问题 $(\widetilde{P}_N)$ 的费用泛函可作类似处理. 事实上, 当 $\vartheta(x,v,\cdot)$ 是几乎处处连续时, 这是当 $N\to\infty$ 时 (6.6) 中的连续性均模 $\tau(\vartheta;h_N)\to 0$ 这一事实的一个简单推论. 若把离散逼近问题 (P_N) 中的费用泛函 (6.20) 替换为简化形式 (6.28), 则所得问题记为 $(\overline{P}_N)$. 下面对 (P_N) 和 $(\overline{P}_N)$ 都予以考虑, 并利用它们来导出原问题的必要最优性条件. 由此所得的结果对初始数据所作假设是不同的, 而这些假设可以用来验证所需的必要最优性条件. 具体地说, 简化问题 $(\overline{P}_N)$ (当 $N\to\infty$ 时) 需要对 ϑ 关于 t 的几乎处处连续性 (而非可测性) 假设, 但另一方面, 它放松了在情形 (P_N) 时对状态空间 X 的要求.

6.1.4　离散时间包含的必要最优性条件

关于离散逼近强收敛的定理 6.13 在离散时间问题 (P_N), 以及注 6.15 中其简化形式 $(\overline{P}_N)$ 的最优解和初始连续时间问题 (P) 的给定松弛中间局部极小点 $\bar{x}(\cdot)$ 之间架起了一座桥梁. 进一步的策略如下: 首先对离散逼近问题序列 (P_N) 和 $(\overline{P}_N)$ 建立必要最优性条件. 然后, 当 $N\to\infty$ 时取极限得到由微分包含控制的初始最优控制问题 (P) 的给定局部极小点的必要条件.

这一小节致力于导出一般离散时间 Bolza 问题及其相应的特殊离散逼近问题 (P_N) 和 $(\overline{P}_N)$ 的必要最优性条件. 这里探究两种方法. 第一种方法涉及将一般的离散时间包含的动态最优化问题化为具有算子约束的数学规划的非动态问题, 然后对这样的问题利用 5.1.2 小节得到的必要最优性条件. 第二种方法是基于离散逼近问题 (P_N) 和 $(\overline{P}_N)$ 的具体特性并且利用 2~4 章中的模糊微分法则. 利用这两种方法导出的结果互不包含, 它们要求不同的假设条件. 然而, 通过第二种方法得到的近似必要最优性条件在下面的小节中更适合导出连续时间问题 (P) 的相应结果, 但是通过第一种方法得到的结果当然也有其独立的意义.

下面从第一种方法开始. 考虑下面具有算子、不等式和几何约束的数学规划

(非动态) 问题 (MP), 可将动态最优化的离散时间问题化为此类问题:

$$\begin{cases} \min & \varphi_0(z) \\ \text{s.t.} & \varphi_j(z) \leqslant 0, \quad j=1,\cdots,s, \\ & f(z)=0, \\ & z \in \Xi_j \subset Z, \quad j=1,\cdots,l, \end{cases} \tag{6.29}$$

其中 φ_j 是 Z 上的实值函数, $f\colon Z \to E$ 是 Banach 空间之间的一个映射并且 $\Xi_j \subset Z$. 这是一个具有算子约束类型的问题, 这在 5.1.2 小节末考虑过, 唯一不同的是, 现在有由集合 Ξ_j 给出的许多几何约束. 正如下面所看到的, (6.29) 中的几何约束产生于离散化的微分包含 (6.3) 并且其数目随着 $N \to \infty$ 时增加. 注意到, 问题 (MP) 是内在非光滑的, 既便在 (6.29) 中的 f 与 φ_j 和生成动态问题的数据是光滑的情形亦如此. 事实上, 非光滑性来源于 (6.29) 中的几何约束, 它反映了原问题 (P) 及其离散逼近问题中由微分和有限差分包含所控制的动态.

为了导出问题 (MP) 的必要最优性条件, 可以利用推论 5.18, 它涉及像 (6.29) 这样的问题, 只是这里具有许多几何约束. 记

$$\Xi := \Xi_1 \cap \cdots \cap \Xi_l,$$

并且将 (6.29) 中的几何约束替换为 $z \in \Xi$. 现在利用推论 5.18 的结果, 则需要进一步通过初始数据给出问题 (MP) 的必要最优性条件. 这可利用第 3 章发展的广义法向量的微分法则和交集的 SNC 性质办到.

命题 6.16(具有许多几何约束的数学规划的必要条件) 设 $\bar{z}$ 为问题 (6.29) 的一个局部最优解, 其中空间 Z 和 E 是 Asplund 空间并且集合 Ξ_j 在 $\bar{z}$ 附近是局部闭的. 假设所有的 φ_i 在 $\bar{z}$ 附近是 Lipschitz 连续的, f 在 $\bar{z}$ 是广义 Fredholm 的且每个 Ξ_j 在该点是 SNC 的. 那么存在实数 $\{\mu_j \in \mathbb{R} \mid j=0,\cdots,s\}$ 以及线性泛函 $e^* \in E^*$ 和 $\{z_j^* \in Z^* \mid j=1,\cdots,l\}$, 不都为零, 使得 $\mu_j \geqslant 0$ $(j=0,\cdots,s)$ 和

$$\mu_j \varphi_j(\bar{z}) = 0, \quad j=1,\cdots,s, \tag{6.30}$$

$$z_j^* \in N(\bar{z}; \Xi_j), \quad j=1,\cdots,l, \tag{6.31}$$

$$-\sum_{j=1}^{l} z_j^* \in \partial\left(\sum_{j=0}^{s} \mu_j \varphi_j\right)(\bar{z}) + D_N^* f(\bar{z})(e^*). \tag{6.32}$$

证明 对由集合 Ξ_j 的交所给出的压缩几何约束 $z \in \Xi$ 下的问题 (6.29), 应用推论 5.18, 则可找到 $\{\mu_j \geqslant 0 \mid j=0,\cdots,s\}$ 和 $e^* \in E^*$, 不都为零, 使得 μ_j 满足 (6.30) 中的互补松弛条件和

$$0 \in \partial\left(\sum_{j=0}^{s} \mu_j \varphi_j\right)(\bar{z}) + D_N^* f(\bar{z})(e^*) + N(\bar{z}; \Xi), \tag{6.33}$$

其中需要交集 Ξ 在 $\bar{z}$ 处是 SNC 的. 由推论 3.81, 如果每个 Ξ_j 在该点是 SNC 的且满足规范条件

$$\left[z_1^* + \cdots + z_l^* = 0, \quad z_j^* \in N(\bar{z}; \Xi_j)\right] \Longrightarrow \left[z_j^* = 0,\ j = 1, \cdots, s\right],$$

那么 Ξ 在 $\bar{z}$ 处是 SNC 的. 更进一步, 当除一个之外所有的 Ξ_j 在 $\bar{z}$ 点是 SNC 的时, 由推论 3.37, 同样的规范条件确保了交集公式

$$N(\bar{z}; \Xi) \subset N(\bar{z}; \Xi_1) + \cdots + N(\bar{z}; \Xi_l).$$

将此替换到 (6.33) 中, 可断定上述规范条件蕴涵 (6.32) 并且 $(\mu_j, e^*) \neq 0$. 同时, 违反规范条件意味着 (6.32) 也成立, 此时 $(z_1^*, \cdots, z_l^*) \neq 0$ 以及所有 μ_j 和 e^* 都为零. 证毕. △

现在考虑将命题 6.16 应用到下面约束的离散时间包含的 Bolza 问题中, 记为 (DP):

$$\begin{cases} \min \quad \varphi(x_0, x_N) + h\displaystyle\sum_{j=0}^{N-1} \vartheta_j\left(x_j, \frac{x_{j+1} - x_j}{h}\right) \\ \text{s.t.} \quad x_{j+1} \in x_j + hF_j(x_j), \quad j = 0, \cdots, N-1, \\ \qquad\ (x_0, x_N) \in \Xi \subset X^2, \end{cases}$$

其中 $F_j \colon X \rightrightarrows X$, φ 和 ϑ_j 是 X^2 上的实值函数且 $h > 0$ 和 $N \in \mathbb{N}$ 固定. 注意到, 问题 (DP) 对任意固定的 N 集成了前面小节中离散逼近问题的基本结构, 但是没有考虑其中与逼近初始连续时间问题 (P) 的给定中间局部极小点 $\bar{x}(\cdot)$ 相关联的项.

定理 6.17(离散时间包含的必要最优性条件) 设 $\{\bar{x}_j \mid j = 0, \cdots, N\}$ 为问题 (DP) 的一个最优解. 假设 X 是 Asplund 的, 集合 Ξ 和 $\operatorname{gph} F_j$ 在点 $(\bar{x}_0, \bar{x}_N)$ 和 $(\bar{x}_j, (\bar{x}_{j+1} - \bar{x}_j)/h)$ 分别是 SNC 的, 并且对所有 $j = 0, \cdots, N-1$, 函数 φ 和 ϑ_j 在相应的点 $\bar{x}_j$ 附近是局部 Lipschitz 的. 那么存在 $\lambda \geqslant 0$ 和 $\{p_j \in X^* \mid j = 0, \cdots, N\}$, 不同时为零, 使得对所有 $j = 0, \cdots, N-1$, 有广义 Euler-Lagrange 包含

$$\left(\frac{p_{j+1} - p_j}{h}, p_{j+1}\right) \in \lambda\partial\vartheta_j\left(\bar{x}_j, \frac{\bar{x}_{j+1} - \bar{x}_j}{h}\right) + N\left(\left(\bar{x}_j, \frac{\bar{x}_{j+1} - \bar{x}_j}{h}\right); \operatorname{gph} F_j\right)$$

和横截性包含

$$(p_0, -p_N) \in \lambda\partial\varphi(\bar{x}_0, \bar{x}_N) + N((\bar{x}_0, \bar{x}_N); \Xi).$$

证明 容易看到, 离散动态最优化问题 (DP) 可等价地写为 (6.29) 中的非动态数学规划形式, 其中

$$z:=(x_0,\cdots,x_N,v_0,\cdots,v_{N-1})\in Z:=X^{2N+1},\quad E:=X^N,\quad l:=N+1,$$
$$\varphi_0(z):=\varphi(x_0,x_N)+h\sum_{j=0}^{N-1}\vartheta_j(x_j,v_j),\quad \varphi_j(z):=0\quad(j\geqslant 1),$$
$$f(z)=(f_0(z),\cdots,f_{N-1}(z)),$$
$$f_j(z):=x_{j+1}-x_j-hv_j,\quad j=0,\cdots,N-1,$$
$$\Xi_j:=\{z\in X^{2N+1}\mid v_j\in F_j(x_j)\},\quad j=0,\cdots,N-1,$$
$$\Xi_N:=\{z\in x^{2N+1}\mid (x_0,x_N)\in\Xi\}.$$

则 $\bar z:=(\bar x_0,\cdots,\bar x_N,(\bar x_1-\bar x_0)/h,\cdots,(\bar x_N-\bar x_{N-1})/h)$ 是具有上述定义数据的数学规划问题 (6.29) 的一个局部最优解. 算子约束映射 f 在 $\bar z$ 点的确是广义 Fredholm 的; 此外, 在关于 F_j 和 Ξ 的假设下, 集合 $\Xi_j,\,j=0,\cdots,N$ 在 $\bar z$ 点显然是 SNC 的. 因为费用函数 φ_0 在 $\bar z$ 附近是局部 Lipschitz 的且 Z 和 E 的乘积空间是 Asplund 的, 将命题 6.16 中的必要最优性条件应用到 (MP) 问题上, 就可得到数 $\mu_0\geqslant 0$ 以及线性泛函 $z_j^*=(x_{0j}^*,\cdots,x_{Nj}^*,v_{0j}^*,\cdots,v_{(N-1)j}^*)\in(X^*)^{2N+1}$ $(j=0,\cdots,N)$ 和 $e^*=(e_0^*,\cdots,e_{N-1}^*)\in(X^*)^N$, 不都为零, 使得条件 (6.30)~(6.32) 对于上述所定义的数据都成立. 由上面关于 Ξ_j 的构造可得 (6.31) 等价于

$$\begin{cases}(x_{jj}^*,v_{jj}^*)\in N\left(\left(\bar x_j,\dfrac{\bar x_{j+1}-\bar x_j}{h}\right);\mathrm{gph}\,F_j\right),x_{ij}^*=v_{ij}^*=0,\\ \qquad\qquad\text{如果}i\neq j,\ j=0,\cdots,N-1;\\ (x_{0N}^*,x_{NN}^*)\in N((\bar x_0,\bar x_N);\Xi)\quad x_{iN}^*=v_{iN}^*=0,\\ \qquad\qquad\text{其他情形}.\end{cases}$$

记 $\lambda:=\mu_0$ 并利用定理 3.36 中局部 Lipschitz 函数次梯度的和法则, 由 (6.32) 以及 φ_0 和 f 的构造可得

$$(x_0^*,x_N^*)\in\partial\varphi(\bar x_0,\bar x_N),\quad (u_j^*,w_j^*)\in\partial\vartheta_j\left(\bar x_j,\frac{\bar x_{j+1}-\bar x_j}{h}\right)$$

对 $j=0,\cdots,N-1$ 满足关系式

$$\begin{cases}-x_{00}^*-x_{0N}^*=\lambda(x_0^*+hu_0^*)-e_0^*,\\ -x_{jj}^*=\lambda hu_j^*+e_{j-1}^*-e_j^*,\quad j=1,\cdots,N-1,\\ -x_{NN}^*=\lambda x_N^*+e_{N-1}^*,\\ -v_{jj}^*=h(\lambda w_j^*-e_j^*),\quad j=0,\cdots,N-1.\end{cases}$$

最后记

$$p_0:=-x_{0N}^*-\lambda x_0^*+e_0^*;\quad p_j:=he_{j-1}^*,\quad j=1,\cdots,N,$$

即可得所要的 Euler-Lagrange 和横截性包含并且 $\lambda\geqslant 0$ 和 $\{p_j\in X^*\mid j=0,\cdots,N\}$ 不同时为零. 证毕. △

回到离散逼近问题 (P_N) 和 $(\overline{P}_N)$ 上来. 固定任意的 $N \in \mathbb{N}$, 注意到由 (6.3), (6.21)~(6.23) 和 (6.28) 定义的问题 $(\overline{P}_N)$ 可化为数学规划 (6.29) 的形式, 这和 (DP) 的形式仅有细微的差别. 事实上, 令

$$z := (x_0, \cdots, x_N, v_0, \cdots, v_{N-1}) \in Z := X^{2N+1}, \quad E := X^N, \quad s := N+2, \quad l := N,$$

则可将 $(\overline{P}_N)$ 重写为 (6.29) 的形式, 其中

$$\varphi_0(z) := \varphi(x_0, x_N) + \|x_0 - \bar{x}(a)\|^2 + h_N \sum_{j=0}^{N-1} \vartheta_j(x_j, v_j) + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \|v_j - \dot{\bar{x}}(t)\|^2 \mathrm{d}t, \tag{6.34}$$

$$\varphi_j(z) := \begin{cases} \|x_{j-1} - \bar{x}(t_{j-1})\| - \varepsilon/2, & j = 1, \cdots, N+1, \\ \displaystyle\sum_{i=0}^{N-1} \int_{t_i}^{t_{i+1}} \|v_i - \dot{\bar{x}}(t)\|^2 \mathrm{d}t - \varepsilon/2, & j = N+2, \end{cases} \tag{6.35}$$

$$\begin{aligned} & f(z) = (f_0(z), \cdots, f_{N-1}(z)), \\ & f_j(z) := x_{j+1} - x_j - h_N v_j, \quad j = 0, \cdots, N-1, \end{aligned} \tag{6.36}$$

$$\begin{aligned} & \Xi_j := \{z \in X^{2N+1} \mid v_j \in F_j(x_j)\}, \quad j = 0, \cdots, N-1, \\ & \Xi_N := \{z \in X^{2N+1} \mid (x_0, x_N) \in \Omega_N\}, \end{aligned} \tag{6.37}$$

其中 $\vartheta_j(x,v) := \vartheta(x,v,t_j)$, $F_j(x) := F(x,t_j)$, $\Omega_N := \Omega + \eta_N \mathbb{B}$. 注意到问题 (DP) 和 $(\overline{P}_N)$ 的 (MP) 形式的唯一不同反映在费用函数和不等式约束的项上, 它们涉及初始连续时间问题 (P) 的给定中间局部极小点 $\bar{x}(\cdot)$. 在导出必要最优性条件的过程中, 类似定理 6.17 的证明, 这些项很容易处理. 此外, 在当 $N \to \infty$ 取极限的过程中, 这些项对必要最优性条件的影响将消失, 即在原问题 (P) 中, 从必要最优性条件的角度来讲, 它们实际上可被忽略; 见下述.

类似地, 注意到 (6.3), (6.20)~(6.23) 定义的问题 (P_N) 可等价地化为 (MP) 形式 (6.29), 它具有费用函数

$$\varphi_0(z) := \varphi(x_0, x_N) + \|x_0 - \bar{x}(a)\|^2 + \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \left[\vartheta(x_j, v_j, t) + \|v_j - \dot{\bar{x}}(t)\|^2\right] \mathrm{d}t \tag{6.38}$$

和相同的约束 (6.35)~(6.37). (6.34) 和 (6.38) 的不同在于将

$$h_N \sum_{j=0}^{N-1} \vartheta_j(x_j, v_j) \quad \text{替换为} \quad \sum_{j=0}^{N-1} \int_{t_j}^{t_{j+1}} \vartheta(x_j, v_j, t) \mathrm{d}t,$$

后者允许处理可和 (在 Bochner 意义下) 被积函数 $\vartheta(x,v,\cdot)$. 为了导出涉及可测/可和被积函数问题的必要最优性条件, 需要一个辅助结果 (当然就其自身而言也是重要的) 来确保可在积分符号下取次微分, 这可看成是 3.2.1 小节得到的关于 Lipschitz 函数的有限和的次微分和法则的一个 "无穷和" (连续测度) 推广. 然而, 该积分结果的正确性要求对所论空间作更多的限制: 相比于定理 6.17 中用到的在有限和法则中的 Asplund 结构, 这里需要假设空间的自反性和可分性. 虽然这个次微分公式在相当广泛的可测空间中也成立, 下面只给出实区间的情形, 即 $T=[0,1]$, 这将在随后的应用中用到. 集值映射的积分总是理解为其可和选择的积分的集合.

引理 6.18(积分泛函的基本次梯度) 设 X 为自反和可分 Banach 空间. 给定 $\bar{x}\in X$, 假设 $\varphi\colon X\times[0,1]\to\mathbb{R}$ 对 $\bar{x}$ 附近的每一个 x 关于 t 是可测的且在 $\bar{x}$ 附近是局部 Lipschitz 的, 在 $[0,1]$ 上具有可和模. 那么有

$$\partial\left(\int_0^1\varphi(\cdot,t)\mathrm{d}t\right)(\bar{x})\subset\mathrm{cl}\int_0^1\partial\varphi(\bar{x},t)\mathrm{d}t, \tag{6.39}$$

其中右端的次微分是关于 x 取的并且闭包 "cl" 是关于 X^* 上的范数拓扑取的.

证明 首先注意到, 映射 $\partial\varphi(\bar{x},\cdot)\colon[0,1]\rightrightarrows X^*$ 是闭值的且在标准的意义下是可测的 (集值映射 $F\colon T\rightrightarrows Y$ 称为可测的, 如果对任何开子集 $\Theta\subset Y$, 逆像 $F^{-1}(\Theta)$ 是可测的); 对于闭值映射, 这样的可测性有许多其他等价的描述; 参见例如 Rockafellar 和 Wets[1165] 中的定理 14.3 和 14.56, 它们在无穷维空间中也成立. 也注意到, 对可分像空间的情形, 该可测性等价于强可测性 (即可由一列阶梯映射几乎处处逐点逼近), 强可测性在所考虑的 Bochner 积分中是明确要求的. 由众所周知的关于可测选择的定理 (参见例如前面提到的书 [1165] 以及 Castaing 与 Valadier 早期的书 [229]), 存在可测单值映射 $\xi\colon[0,1]\to X^*$, 使得 $\xi(t)\in\partial\varphi(\bar{x},t)$ 对几乎处处 $t\in[0,1]$ 成立. 此外, 因为 X^* 是可分的且 $\partial\varphi(\bar{x},\cdot)$ 是可积有界的, 它由 $\varphi(\cdot,t)$ 的可和 Lipschitz 模界定 (这容易从所作的假设中得到; 参见引理 1.81), 所以 $\partial\varphi(\bar{x},\cdot)$ 的每个可测选择 ξ 在 $[0,1]$ 上是 Bochner 可积的. 因此, (6.39) 右端的多值积分是良好定义的并且是非空的.

由 Clarke[225, 定理 2.7.2] 可知, 如果 (6.39) 两端的 Lipschitz 函数的基本次微分用其 Clarke 广义梯度来替换, 那么也成立. 现在利用定理 3.57 和 X 的自反性有

$$\partial\left(\int_0^1\varphi(\cdot,t)\mathrm{d}t\right)(\bar{x})\subset\int_0^1\mathrm{cl}\,\mathrm{co}\,\partial\varphi(\bar{x},t)\mathrm{d}t,$$

这是因为由 Mazur 定理, 在自反空间中弱闭包和范数闭包是一致的. 另一方面, 对每个紧值、强可测和可积有界映射 F, 由著名的 Lyapunov-Aumann 定理的无穷维推广 (参见例如 Tolstonogov [1258] 中的 1.1 节) 可知

$$\int_0^1\mathrm{cl}\,\mathrm{co}\,F(t)\mathrm{d}t=\mathrm{cl}\int_0^1F(t)\mathrm{d}t.$$

这就得到 (6.39). 证毕. △

根据定理 6.17 和随后的讨论, 类似地可以得到并且证明离散逼近问题 (P_N) 和 $(\overline{P}_N)$ 最优解的广义 Euler-Lagrange 和横截性包含. 定理 6.17 中问题 (DP) 和 $(\overline{P}_N)$ 的必要最优性条的区别仅在于当 $N \to \infty$ 时收敛于零的项不同. 问题 (P_N) 的 Euler-Lagrange 包含平行于 $(\overline{P}_N)$ 中的 Euler-Lagrange 包含, 只是右端的

$$\lambda_N \partial\vartheta\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t_j\right)$$

被替换为

$$\frac{\lambda_N}{h_N}\int_{t_j}^{t_{j+1}} \partial\vartheta\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t\right) \mathrm{d}t$$

的范数闭包, 该范数闭包来源于引理 6.18 中的积分公式. 当 $N \to \infty$ 时, 对几乎处处 $t \in [a, b]$, 后者收敛于 $\lambda\partial\vartheta(\bar{x}(t), \dot{\bar{x}}(t), t)$; 参见定理 6.21 的证明.

由这种方法得到的结果利用了定理 6.16 中关于一般数学规划问题的确切/极限最优性条件, 这在问题 (P_N) 和 $(\overline{P}_N)$ 中要求集合 $\operatorname{gph} F_j$ 和 Ω_N 的 SNC 性质. 这些假设对导出初始连续时间问题 (P) 的必要最优性条件的极限过程可能有所限制; 因此, 下面将试图避免或实质性地放宽这些假设. 从数学规划问题的近似/模糊必要最优性条件开始, 并大量利用离散时间问题 (P_N) 和 $(\overline{P}_N)$ 的具体特性, 这是可以办到的. 为了实现这个方法, 在问题 (DP) 中需要利用集值映射的类 Lipschitz 性质. 该集值映射 F_j 生成了图像的几何约束, 在 (P_N) 和 $(\overline{P}_N)$ 中亦然, 定理 6.17 中并没有作这样的假设. 另一方面, (6.1) 中初始映射 $F(\cdot, t)$ 的 Lipschitz 连续性在假设之中 (参见 6.1.1 小节中的 (H1)). 因此, 在必要最优性条件离散逼近的上下文中, 可以毫无顾虑地利用它.

下面的两个定理在离散时间问题序列 $(\overline{P}_N)$ 和 (P_N) 中给出了局部极小点的近似必要最优性条件. 它们的证明涉及利用前面章节中的一些模糊/邻域分析法则. 尤其是, 利用定理 2.33 中 Fréchet次梯度的半 -Lipschitz 和法则和引理 3.1 中 Fréchet法向量的模糊交集法则. 这些结果是通过那些任意靠近参考点处的点的 Fréchet次梯度和法向量给出了在参考点处的和或交集的对应的次梯度或法向量. 仅仅是为了记号的简便, 在下面定理的公式和证明中假设问题中的任意靠近的点可归结为参考点. 本节的主要目的是导出连续时间问题 (P) 的必要最优性条件, 这将在下一小节中完成, 从这个角度来看, 此约定实际上并没有限制一般性. 事实上, 在提到的点之间的可能不同之处在取极限的过程中显然会消失. 有兴趣的读者可自行补上所有细节.

下面从注 6.15 描述的简化的离散逼近问题 $(\overline{P}_N)$ 当 $N \to \infty$ 时的近似必要最优条件开始. 若在原问题 (P) 中的被积项 $\vartheta(x, v, \cdot)$ 假设为几乎处处连续的, 则这些

条件对 (P) 也有效. 下面依惯例, $\mathbb{B}^*$ 表示对偶闭球 (不再特别指明空间) 并且 ϑ 的次微分是关于前两个变量取的.

定理 6.19(简化离散时间问题的近似 Euler-Lagrange 条件) 设 $\bar{x}_N(\cdot) = \{\bar{x}_N(t_j) \mid j = 0, \cdots, N\}$ 为问题 $(\overline{P}_N)$ 当 $N \to \infty$ 时的局部最优解. 假设 X 是 Asplund 的, Ω_N 在 $(\bar{x}_N(t_0), \bar{x}_N(t_N))$ 附近是局部闭的, F_j 在 $(\bar{x}_N(t_j), [\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)]/h_N)$ 附近是闭图和类 Lipschitz 的, 并且函数 φ 和 $\vartheta(\cdot, \cdot, t_j)$ 对每个 $j = 0, \cdots, N-1$ 是局部 Lipschitz 的. 考虑

$$\theta_{Nj} := 2\int_{t_j}^{t_{j+1}} \left\| \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t) \right\| \mathrm{d}t, \quad j = 0, \cdots, N-1.$$

那么存在一个不依赖 N 的数 $\gamma > 0$, 使得对某自然数序列 $N \to \infty$ 和正数序列 $\varepsilon_N \downarrow 0$, 有乘子 $\lambda_N \geqslant 0$ 和伴随轨道 $p_N(\cdot) = \{p_N(t_j) \in X^* \mid j = 0, \cdots, N\}$ 满足非平凡条件

$$\lambda_N + \|p_N(t_N)\| \geqslant \gamma \quad (N \to \infty), \tag{6.40}$$

近似 Euler-Lagrange 包含

$$\begin{aligned} &\left(\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N}, p_N(t_{j+1}) - \lambda_N \frac{\theta_{Nj}}{h_N} b_{Nj}^* \right) \\ &\in \lambda_N \widehat{\partial}\vartheta \left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t_j \right) \\ &\quad + \widehat{N}\left(\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} \right); \operatorname{gph} F_j \right) + \varepsilon_N \mathbb{B}^*, \end{aligned} \tag{6.41}$$

其中 $j = 0, \cdots, N-1$, 并且近似横截性包含

$$\begin{aligned} &(p_N(t_0) - 2\lambda_N b_N^* \|\bar{x}(a) - \bar{x}_N(t_0)\|, -p_N(t_N)) \\ &\in \lambda_N \widehat{\partial}\varphi(\bar{x}_N(t_0), \bar{x}_N(t_N)) + \widehat{N}((\bar{x}_N(t_0), \bar{x}_N(t_N)); \Omega_N) + \varepsilon_N h_N \mathbb{B}^* \end{aligned} \tag{6.42}$$

对某 $b_N^*, b_{Nj}^* \in \mathbb{B}^*$ 成立.

证明 对固定的 $N \in \mathbb{N}$, 考虑问题 $(\overline{P}_N)$ 的 (MP) 等价形式 (6.29), 其中数据在 (6.34)~(6.37) 中定义. 记

$$\bar{z} := (\bar{x}_N(t_0), \cdots, \bar{x}_N(t_N), \bar{v}_N(t_0), \cdots, \bar{v}_N(t_{N-1})),$$

并且取足够大的 N 使得关于 $\bar{x}_N(\cdot)$ 的约束 (6.22) 和 (6.23) 以严格不等式形式成立. 这可由定理 6.13 的强收敛结果容易得到.

首先假设 (6.36) 中的 f 在 $\bar{z}$ 点关于交集 $\varXi := \varXi_0 \cap \cdots \cap \varXi_N$ 是度量正则的, 其中集合 $\varXi_j$ 已在 (6.37) 中构造. 因为 (6.34) 中的 φ_0 在 $\bar{z}$ 点附近是局部 Lipachitz 的且由 N 的选取, 利用定理 5.16, 可找到 $\mu > 0$ 使得 $\bar{z}$ 是无约束问题

$$\min \quad \varphi_0(z) + \mu(\|f(z)\| + \operatorname{dist}(z; \varXi))$$

的一个局部最优解. 于是, 由广义 Fermat 法则, 有

$$0 \in \widehat{\partial}\big(\varphi_0(\cdot) + \mu\|f(\cdot)\| + \mu\,\mathrm{dist}(\cdot;\varXi)\big)(\bar{z}).$$

现在利用定理 2.33 中的模糊和法则并利用记号约定, 固定任何 $\varepsilon > 0$ 可得

$$0 \in \widehat{\partial}\varphi_0(\bar{z}) + \mu\widehat{\partial}\|f(\cdot)\|(\bar{z}) + \mu\widehat{\partial}\,\mathrm{dist}(\bar{z};\varXi) + (\varepsilon/3)\mathbb{B}^*.$$

由关于距离函数 Fréchet次梯度的命题 1.95 以及 (6.36) 中光滑映射 f 的复合 $\|f(z)\| = (\psi \circ f)(z)$ 的基本链式法, 其中 $\psi(y) := \|y\|$, 则有

$$0 \in \widehat{\partial}\varphi_0(\bar{z}) + \sum_{j=0}^{N-1} \nabla f_j(\bar{z})^* e_j^* + \widehat{N}(\bar{z};\varXi) + (\varepsilon/3)\mathbb{B}^*,$$

其中 $e_j^* \in X^*$. 由 (6.36) 中 $f(z)$ 的结构注意到

$$\sum_{j=0}^{N-1} \nabla f_j(\bar{z})^* e_j^* = (-e_0^*, e_0^* - e_1^*, \cdots, e_{N-2}^* - e_{N-1}^*, e_{N-1}^*, -h_N e_0^*, \cdots, -h_N e_{N-1}^*).$$

更进一步, 由引理 3.1 中的模糊交法则及紧接其后的讨论, 考虑到记号的约定有

$$\widehat{N}(\bar{z};\varXi) \subset \widehat{N}(\bar{z};\varXi_0) + \cdots + \widehat{N}(\bar{z};\varXi_N) + (\varepsilon/3)\mathbb{B}^*.$$

为了证明该式, 需要对涉及的集合验证模糊规范条件 (3.9). 由 (6.37) 中关于这些集合的结构, 模糊规范条件 (3.9) 对集合 $\varXi_j$, $j = 0, \cdots, N-1$ 的交显然成立. 在验证该条件的最后一步中, 需要证明存在 $\gamma > 0$, 使得

$$\left(\widehat{N}\big(z;\bigcap_{j=0}^{N-1}\varXi_j\big) + \gamma\mathbb{B}^*\right) \cap \Big(-\widehat{N}(z_N;\varXi_N) + \gamma\mathbb{B}^*\Big) \cap \mathbb{B}^* \subset \frac{1}{2}\mathbb{B}^*$$

对任意 $z \in \varXi_j \cap (\bar{z} + \gamma\mathbb{B})$, $j = 0, \cdots, N-1$, 并且 $z_N \in \varXi_N \cap (\bar{z} + \gamma\mathbb{B})$ 成立. 由 (6.37) 中集合的结构直接可得, 对任意 $z_j^* \in \widehat{N}(z_j;\varXi_j)$, 其中 $z_j^* = (x_{0j}^*, \cdots, x_{Nj}^*, v_{0j}^*, \cdots, v_{N-1j}^*)$ 和 $z_j = (x_{0j}, \cdots, x_{Nj}, v_{0j}, \cdots, v_{N-1j})$ 靠近 $\bar{z}$, 有关系式

$$x_{ij}^* \in \widehat{D}^* F_j(x_{jj}, v_{jj})(-v_{jj}^*), \quad x_{ij}^* = v_{ij}^* = 0 \quad (\text{如果} i \neq j),$$

其中 $j = 0, \cdots, N-1$, 以及在其他情形, 有

$$(x_{0N}^*, x_{NN}^*) \in \widehat{N}((x_{0N}, x_{NN});\varOmega_N); \quad \text{并且} x_{iN}^* = v_{iN}^* = 0.$$

于是, 如果 F_j 在 (x_{jj}, v_{jj}) 附近是模为 ℓ 类 Lipschitz 的, 那么由关于 Lipschitz 映射 Fréchet上导数的定理 1.43 得到估计

$$\|x_{jj}^*\| \leqslant \ell\|v_{jj}^*\|, \quad j = 0, \cdots, N-1.$$

考虑到上面的模糊规范条件对前面所有步骤都成立, 其中 $\varepsilon_N := \varepsilon/N$, 最后一步显然蕴涵该条件.

接下来, 着手于 (6.34) 中费用函数 φ_0 的 Fréchet次梯度的估计. 众所周知, 对任意 Banach 空间, 在凸分析中有

$$\partial\|\cdot\|^2(x) \subset 2\|x\|\mathbb{B}^* \quad (\forall x \in X).$$

利用该结果并将定理 2.33 中的模糊和法则应用到 (6.34) 中 φ_0 的具体形式上可得到

$$\begin{aligned}\widehat{\partial}\varphi_0(\bar{z}) \subset\, & \widehat{\partial}\varphi(\bar{x}_N(t_0), \bar{x}_N(t_N)) + 2\|\bar{x}_N(t_0) - \bar{x}(a)\|\mathbb{B}^* \\ & + h_N \sum_{j=0}^{N-1} \Big[\widehat{\partial}\vartheta_j(\bar{x}_N(t_j), \bar{v}_N(t_j)) + (0, 2\theta_{Nj}\mathbb{B}^*)\Big] + (\varepsilon/3)\mathbb{B}^*,\end{aligned}$$

其中考虑到记号的约定和 θ_{Nj} 的构造. 现在在广义 Fermat 法则中将上面的关系和估计结合起来可得到

$$\begin{cases} -x_{00}^* - x_{0N}^* - x_0^* - 2b_N^*\|\bar{x}_N(t_0) - \bar{x}(a)\| - u_0^* + e_0^* \in \varepsilon\mathbb{B}^*, \\ -x_{jj}^* - h_N u_j^* - e_{j-1}^* + e_j^* \in \varepsilon\mathbb{B}^*, \quad j = 0, \cdots, N-1, \\ -x_{NN}^* - x_N^* - e_{N-1}^* \in \varepsilon\mathbb{B}^*, \\ -v_{jj}^* - h_N w_j^* - \theta_{Nj} b_{Nj}^* + h_N e_j^* \in \varepsilon\mathbb{B}^*, \quad j = 0, \cdots, N-1, \end{cases}$$

其中 $b_{Nj}^*, b_N^* \in \mathbb{B}^*$, 以及

$$(x_{ij}^*, v_{ij}^*) \in \widehat{N}\left(\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\right); \operatorname{gph} F_j\right),$$

$$(x_0^*, x_N^*) \in \widehat{\partial}\varphi(\bar{x}_N(t_0), \bar{x}_N(t_N)), \quad (u_j^*, w_j^*) \in \widehat{\partial}\vartheta_j\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\right)$$

对所有 $j = 0, \cdots, N-1$ 成立. 记

$$p_N(t_0) := -x_{0N}^* - \lambda_N x_0^* + e_0^*, \quad p_N(t_j) := e_{j-1}^*, \quad j = 1, \cdots, N,$$

对任意充分大的 $N \in \mathbb{N}$ 和任意 $\varepsilon = \varepsilon_N$, 可得到近似 Euler-Lagrange 和横截性包含 (6.41) 和 (6.42) 且 $\lambda_N = 1$. 注意到, 在所考虑的度量正则性的情形下, 当 $\gamma_N = 1$ 时, 非平凡条件 (6.40) 显然满足.

剩下来考虑 (6.36) 中映射 f 在 $\bar{z}$ 点相对于集合的交 $\varXi := \varXi_0 \cap \cdots \cap \varXi_N$ 不是度量正则的. 对于该情形, 在定义 1.47(ii) 的意义之下, 扩张映射 $f_\varXi(z) := f(z) + \Delta(z; \varXi)$ 在 $\bar{z}$ 点附近不是度量正则的. 现在应用定理 4.5 中得到的 Asplund 空间中度量正

则性的邻域刻画, 不难观察到该准则可等价地写为: Asplund 空间之间的闭图映射 $F\colon X \rightrightarrows Y$ 在 $(\bar{x},\bar{y})\in \operatorname{gph} F$ 点附近是度量正则的, 当且仅当存在一个正数 ν 使得

$$\ker \widehat{D}^*F(x,y)\subset \mathbb{B}^*,\quad \text{只要}x\in \bar{x}+\nu\mathbb{B},\quad y\in F(x)\cap(\bar{y}+\nu\mathbb{B}).$$

将此结果应用到映射 $f(z)+\Delta(z;\Xi)$ 上, 该映射在 $\bar{z}$ 点附近不是度量正则的, 那么当 N 固定时, 有下述断言: 对任何 $\eta>0$, 存在 $z\in\bar{z}+\eta\mathbb{B}$ 和 $e^*\in\ker\widehat{D}^*f_{\Xi}(z)$, 满足 $\|e^*\|>1$, 其中 $e^*=(e_0^*,\cdots,e_{N-1}^*)\in(X^*)^N$. 因此,

$$0\in\widehat{D}^*f_{\Xi}(z)(e^*)\quad \text{对某些 } \|e^*\|>1 \text{ 和 } z\in\bar{z}+\eta\mathbb{B} \text{ 成立}.$$

固定 $\varepsilon>0$, 利用定理 1.62 (i) 中上导数的和法则和上面关于 Fréchet法向量交集法则有

$$0\in\sum_{j=0}^{N-1}\nabla f_j(z)(e_j^*)+\sum_{j=0}^{N}\widehat{N}(z_j;\Xi_j)+\varepsilon\mathbb{B}^*$$

对某些 $z_j\in\Xi_j\cap(z+\varepsilon\mathbb{B})$ 成立. 根据记号的约定, 为了简单起见, 记 $z_j=z=\bar{z}$. 因此, 存在 $z_j^*\in\widehat{N}(\bar{z};\Xi_j)$ 满足

$$-\sum_{j=0}^{N}z_j\in\sum_{j=0}^{N-1}\nabla f_j(\bar{z})^*e_j^*+\varepsilon\mathbb{B}^*.$$

考虑到 (6.36) 中映射 f 的结构和 (6.37) 中的集合 Ξ_j, 可如上找到对偶元

$$(x_{ij}^*,v_{ij}^*)\in\widehat{N}\left(\left(\bar{x}_N(t_j),\frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N}\right);\operatorname{gph}F_j\right),\quad j=0,\cdots,N-1$$

和

$$(x_{0N}^*,x_{NN}^*)\in\widehat{N}((\bar{x}_N(t_0),\bar{x}_N(t_N));\Omega_N),$$

满足关系式

$$\begin{cases}-x_{00}^*-x_{0N}^*+e_0^*\in\varepsilon\mathbb{B}^*,\\ -x_{jj}^*-e_{j-1}^*+e_j^*\in\varepsilon\mathbb{B}^*,\quad j=0,\cdots,N-1,\\ -x_{NN}^*-x_N^*-e_{N-1}^*\in\varepsilon\mathbb{B}^*,\\ -v_{jj}^*+h_Ne_j^*\in\varepsilon\mathbb{B}^*,\quad j=0,\cdots,N-1.\end{cases}$$

定义伴随轨道 $p_N(t_j)$, $j=0,\cdots,N$, 如下

$$p_N(t_0):=-x_{0N}^*+e_0^*;\quad p_N(t_j):=e_{j-1}^*,\quad j=1,\cdots,N.$$

由上面的构造可知, 对 $\lambda_N=0$ 和任意的 $\varepsilon_N=\varepsilon>0$, $(\bar{x}_N(\cdot),p_N(\cdot))$ 满足 Euler-Lagrange 包含 (6.41) 和横截性包含 (6.42). 更进一步, 伴随轨道 $p_N(\cdot)$ 遵循如下非平凡条件:

$$\|p_N(t_1)\|+\cdots+\|p_N(t_N)\|\geqslant 1,\quad \text{对所有大的 } N\in\mathbb{N} \text{ 成立}.$$

最后, 由 F_j 上的类 Lipschitz 假设, 下面证明在该情形下非平凡条件可等价地写为 $\|p_N(t_N)\| \geqslant 1$, 这和 (6.40) 式中 $\lambda_N = 0$ 时的情形一致. 现在近似 Euler-Lagrange 包含 (6.41) 可改写为如下形式

$$\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N} \in \widehat{D}^* F_j \left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N} \right) (-p_N(t_{j+1}) + \mathbb{B}^*) \\ + \varepsilon \mathbb{B}^*, \quad j = 0, \cdots, N-1$$

那么由定理 1.43, 定理中假设 F_j (具有模 $\ell = \ell_F$) 的类 Lipschitz 性质可得到 $\|x_j^*\| \leqslant \ell \|v_j^*\|$ 对任何 $(\bar{x}_N(t_j), [\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)]/h_N)$ 附近的 (x_j, v_j) 和 $x_j^* \in \widehat{D}^* F_j(x_j, v_j)(v_j^*)$ 成立. 因此,

$$\|p_N(t_{N-1})\| \leqslant \|p_N(t_N)\|(1 + h_N \ell) + h_N \varepsilon.$$

继续该过程有

$$\|p_N(t_j)\| \leqslant \exp(\ell(b-a)) \|p_N(t_N)\| + \varepsilon(b-a), \quad j = 0, \cdots, N.$$

在所考虑的情形 $\lambda_N = 0$ 下, 假设与 (6.41) 和 (6.42) 一道的非平凡条件 (6.40) 不成立. 选取序列 $\gamma_k \downarrow 0$ (当 $k \to \infty$ 时) 及常数 N_k 和 ε_k, 使得

$$N_k := [1/\gamma_k], \quad \varepsilon_k \leqslant \gamma_k^2, \quad \|p_N(t_N)\| \leqslant \gamma_k^2 \quad (k \to \infty),$$

其中 $[\cdot]$ 表示小于或等于所给定实数的最大整数. 由伴随轨道的估计有

$$\sum_{j=1}^{N_k} \|p_{N_k}(t_j)\| \leqslant 2N_k \gamma_k \exp(\ell(b-a)) + \varepsilon_k N_k (b-a) \\ \leqslant \gamma_k \exp(\ell(b-a)) + \gamma_k (b-a) \downarrow 0, \quad k \in \mathbb{N},$$

这和上面建立的事实相矛盾. 证毕. △

在本小节最后, 得到由 (6.3), (6.20)~(6.23) 定义的离散问题序列 (P_N) 的近似必要最优性条件. 这些问题和简化问题 $(\overline{P}_N)$ 的不同在于 (P_N) 处理原问题中近似可和被积函数 $\vartheta(x, v, \cdot)$, 这在费用函数 (6.20) 中涉及 ϑ 的被积项中得到反映. 与 $(\overline{P}_N)$ 中的项相比, 该被积项使得对问题 (P_N) 的分析要复杂得多. 为此, 需要利用关于在 Bochner 积分号下取次微分的引理 6.18, 这要求对空间 X 作另外的假设. 下面的定理在 (P_N) 的广义 Euler-Lagrange 包含的框架下将这些进展结合在一起. 这里继续使用定理 6.19 之前讨论过的记号约定.

定理 6.20(涉及可和被积函数离散逼近问题的近似 Euler-Lagrange 条件) 设 $\bar{x}_N(\cdot) = \{\bar{x}_N(t_j) \mid j = 0, \cdots, N\}$ 为问题 (P_N) (当 $N \to \infty$ 时) 的局部最优解. 假

设 X 是自反和可分的, φ, F_j, Ω_N 和 θ_{Nj} 与定理中的一样且 ϑ 满足 6.1.3 小节假设 (H3), 其中的连续性替换为 Lipschitz 连续性. 那么存在一个不依赖 N 的数 $\gamma>0$, 使得对某自然数序列 $N\to\infty$ 和正数序列 $\varepsilon_N\downarrow 0$ 有乘子 $\lambda_N\geqslant 0$ 及伴随轨道 $p_N(\cdot)=\{p_N(t_j)\in X^*\mid j=0,\cdots,N\}$, 满足非平凡条件 (6.40), 近似横截性包含 (6.42) 和修改的 Euler-Lagrange 包含

$$\begin{aligned}&\left(\frac{p_N(t_{j+1})-p_N(t_j)}{h_N},p_N(t_{j+1})-\lambda_N\frac{\theta_{Nj}}{h_N}b_{Nj}^*\right)\\&\in\frac{\lambda_N}{h_N}\operatorname{cl}\int_{t_j}^{t_{j+1}}\partial\vartheta\left(\bar{x}_N(t_j),\frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N},t_j\right)\mathrm{d}t\\&+\widehat{N}\left(\left(\bar{x}_N(t_j),\frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N}\right);\operatorname{gph}F_j\right)+\varepsilon_Nh_N\mathbb{B}^*\end{aligned}\tag{6.43}$$

对所有 $j=0,\cdots,N-1$ 和某 $b_{Nj}^*\in\mathbb{B}^*$ 成立.

证明 每个问题 (P_N) 可等价地写为 (MP) 形式 (6.29) 并且有 (6.35)~(6.38) 中所定义的数据. 余下的论证类似定理 6.19, 另外利用引理 6.18 来计算积分函数的次梯度. 在定理中关于 X 的额外假设下这是可以做到的, 从而得到近似 Euler-Lagrange 包含的修改形式 (6.43). △

考虑到定理 6.14 的值收敛结果, 可将本小节得到的离散逼近问题的必要最优性条件看成原问题 (P) 的次最优性条件. 此外, 由定理 6.13 和注 6.15 给出的强收敛结果, 可将上面离散逼近问题的必要最优性条件视为原问题关于给定松弛中间局部极小点的次最优性条件. 注意到, 定理 6.13 和 6.14 中的假设确保了离散逼近最优解的存在, 但是初始连续问题 (P) 在有限维或无穷维的情形都并非如此. 下面考虑问题 (P) 的松弛局部极小点的必要最优性条件.

6.1.5 松弛极小点的 Euler-Lagrange 条件

本小节的目的是对涉及约束微分包含的原 Bolza 问题 (P) 导出松弛中间局部极小点的必要条件, 这可通过对前一节得到的离散逼近的必要条件取极限来实现. 这基于定理 6.13 中离散逼近强收敛的结果、定理 6.19 和定理 6.20 中离散问题 (P_N) 和 $(\overline{P}_N)$ 的近似必要最优性条件及广义微分结构的稳定性性质. 该极限过程的要素是对伴随轨道建立一个适当的收敛, 从而可在近似 Euler-Lagrange 包含中取极限. 这可通过下面利用关于 Lipschitz 稳定性上导数的刻画来办到. 该刻画在前一小节中也被用过.

首先阐明该小节主要结果所需要的假设. 这些假设当然包括那些确保离散逼近强收敛和下面用到的满足离散时间问题 (P_N) 和 $(\overline{P}_N)$ 的 (近似) 必要最优性条件的假设. 事实上, 对离散逼近取极限来导出原 Bolza 问题 (P) 的逐点必要最优性条件的过程中并不需要添加太多条件.

下面保持 6.1.1 小节关于 (6.1) 中映射 F 的假设 (H1) 和 (H2) 并考虑 6.1.3 小节假设 (H3) 和 (H4) 的关于 Lipschitz 性质的修改:

(H3′) $\vartheta(\cdot,\cdot,t)$ 在 $U\times(m_F\mathbb{B})$ 上关于 $t\in[a,b]$ 是一致 Lipschitz 连续的, 而 $\vartheta(x,v,\cdot)$ 在 $[a,b]$ 上可测且其范数由一个关于 $(x,v)\in U\times(m_F\mathbb{B})$ 一致的可和函数界定.

(H4′) φ 在 $U\times U$ 上是 Lipschitz 连续的; $\Omega\subset X\times X$ 在 $(\bar{x}(a),\bar{x}(b))$ 附近是局部闭的并使得集合 $\operatorname{proj}_1\Omega\cap(\bar{x}(a)+\varepsilon\mathbb{B})$ 对某 $\varepsilon>0$ 是紧的.

注意到 (H3′) 包含了关于 $\vartheta(x,v,\cdot)$ 的可测性的假设, 这对应定理 6.20. 与定理 6.19 相比较, 定理 6.20 在状态空间 X 上作了更多的限制要求, 这和问题 $(\overline{P}_N)$ 收敛结果中的 $\vartheta(x,v,\cdot)$ 几乎处处连续相关联; 参见注 6.15. 由此也考虑 (H3) 的另外一个修改, 它是上面假设 (H3′) 的一个替代:

(H3″) $\vartheta(x,v,\cdot)$ 在 $[a,b]$ 上几乎处处连续且在该区间上关于 $(x,v)\in U\times(m_F)\mathbb{B}$ 一致有界, 而 $\vartheta(\cdot,\cdot,t)$ 在

$$\Theta_\nu(t):=\{(x,v)\in U\times(m_F+\nu)\mathbb{B}\mid\exists\tau\in(t-\nu,t],\quad \text{满足 } v\in F(x,\tau)\}$$

上关于 $t\in[a,b]$ 对于某 $\nu>0$ 是一致 Lipschitz 连续的.

在涉及 t 的极限过程中, 若连续映射 $F(x,\cdot)$ 和 $\vartheta(x,v,\cdot)$ 为几乎处处连续的, 则这里对移动集合 $\operatorname{gph}F(\cdot)$ 应用定义 5.69 中的广义法锥 N_+, 对 $\vartheta(x,v,t)$ 应用相应的广义次微分. 虽然这些构造可能与非自治目标情形的基本法锥和次微分不同, 但是在一般的确保正则半连续的情形, 它们是一致的; 参见定义 5.69 后面的结果和讨论. 若如 (H3′) 中那样假设 ϑ 关于 t 的可测性, 则不需要将被积函数 ϑ 的基本次微分替换为广义次微分. 在下面的 6.1.6 小节中, 处理微分包含中可测集值映射时, 也不需要替换 $\operatorname{gph}F$ 的基本法锥.

给定满足 $\bar{v}\in F(\bar{x},\bar{t})$ 的 $(\bar{x},\bar{v},\bar{t})$, 对 Asplund 空间中闭子集的情形, 移动集合 $\operatorname{gph}F(t)$ 在点 $(\bar{x},\bar{v})\in\operatorname{gph}F(\bar{t})$ 的广义法锥为

$$N_+\big((\bar{x},\bar{v});\operatorname{gph}F(\bar{t})\big):=\mathop{\operatorname{Limsup}}_{(x,v,t)\to(\bar{x},\bar{v},\bar{t})}\widehat{N}\big((x,v);\operatorname{gph}F(t)\big).$$

相应地, $\vartheta(\cdot,\cdot,\bar{t})$ 在点 $(\bar{x},\bar{v})$ 的广义次微分为

$$\partial_+\vartheta(\bar{x},\bar{v},\bar{t}):=\mathop{\operatorname{Limsup}}_{(x,v,t)\to(\bar{x},\bar{v},\bar{t})}\widehat{\partial}\vartheta(x,v,t),$$

其中 $\widehat{\partial}\vartheta(\cdot,\cdot,t)$ 是对固定的 t 关于 (x,v) 取的. 注意到, $\partial_+\vartheta(\bar{x},\bar{v},\bar{t})$ 可通过移动上图集合 $\operatorname{epi}\vartheta(t)$ 的广义法锥 N_+ 等价描述. 可以看到, 当 F 和 ϑ 不依赖于 t 以及在上面讨论的更一般的情形时, 这些广义结构化为基本结构 $N(\cdot;\operatorname{gph}F)$ 和 $\partial\vartheta$.

现在可以来系统地阐述和证明原 Bolza 问题 (P) 中松弛中间极小点的广义 Euler-Lagrange 条件. 考虑两种情形: 当被积函数 ϑ 关于 t 是几乎处处连续和可积时. 虽然第二种情形对被积函数的要求少且有 Euler-Lagrange 包含的较好形式, 但是第一种情形能够在更一般的 Banach 空间中得到必要最优性条件. 下面从第一种情形开始, 用到的强 PSNC 性质在 3.1.1 小节中已定义并且讨论过.

定理 6.21(具有几乎处处连续被积函数的 Bolza 问题松弛局部极小点的广义 Euler-Lagrange 条件) 设 $\bar{x}(\cdot)$ 为 Bolza 问题 (P) 在假设 (H1), (H2), (H4′) 和 (H3″) 下的松弛中间局部极小点. 假设空间 X 和 X^* 都是 Asplund 的且集合 Ω 在 $(\bar{x}(a), \bar{x}(b))$ 点关于第二个分量是 PSNC 的. 那么存在 $\lambda \geqslant 0$ 和绝对连续映射 $p\colon [a,b] \to X^*$, 不都为零, 满足广义 Euler-Lagrange 包含

$$\begin{aligned}\dot{p}(t) \in \operatorname{cl}\operatorname{co}\Big\{u \in X^* \Big| &(u, p(t)) \in \lambda\partial_+\vartheta(\bar{x}(t), \dot{\bar{x}}(t), t)\\ &+N_+((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F(t))\Big\}, \quad \text{a.e. } t \in [a,b]\end{aligned} \tag{6.44}$$

和横截性包含

$$(p(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega). \tag{6.45}$$

证明 在定理 6.19 中, 考虑到简化离散逼近的强收敛 (参见定理 6.13 和注 6.15), 在离散时间问题 $(\overline{P}_N)$ 的必要最优性条件中取极限即可导出这些条件. 事实上, X 的 Asplund 性质等价于 X^* 的 Radon-Nikodým性质; 参见 6.1.1 小节. 因为 X 是 X^{**} 的闭子空间且 X^* 是 Asplund 的, 可以得到 X 具有 Radon-Nikodým性质. 因此, 定理 6.13 的所有假设条件都满足, 于是可利用离散逼近的强收敛.

注意到, 所作的假设显然满足定理 6.19 的假设条件. 利用那里得到的关于 $(\overline{P}_N)$ 的必要最优性条件, 可以找到 (子) 序列 $\lambda_N \geqslant 0$ 和离散伴随轨道列 $p_N(\cdot) = \{p_N(t_j) \mid j = 0, \cdots, N\}$, 对某 $\varepsilon_N \downarrow 0$ (当 $N \to \infty$ 时) 满足包含 (6.40)~(6.42). 不失一般性, 非平凡条件 (6.40) 可等价地写为

$$\lambda_N + \|p_N(t_N)\| = 1 \quad (\forall N \in \mathbb{N}),$$

这是因为 $\gamma > 0$ 不依赖于 N. 也总是可以假设当 $N \to \infty$ 时, $\lambda_N \to \lambda \geqslant 0$.

下面利用记号 $\bar{x}_N(t)$ 和 $p_N(t)$ 表示相应的离散轨道在 $[a,b]$ 上的分段线性扩张, 它们具有分段常数导数 $\dot{\bar{x}}_N(t)$ 和 $\dot{p}_N(t)$. 对在定理 6.19 中定义的 θ_{Nj}, 考虑由

$$\theta_N(t) := \frac{\theta_{Nj}}{h_N} b^*_{Nj}, \quad 对\ t \in [t_j, t_{j+1}), \quad j = 0, \cdots, N-1$$

给出的一列函数 $\theta_N\colon [a,b] \to \mathbb{R}$. 由定理 6.13 可得

$$\int_a^b \|\theta_N(t)\|\mathrm{d}t \leqslant \sum_{j=0}^{N-1} \theta_{Nj} \leqslant 2\sum_{j=0}^{N-1}\int_{t_j}^{t_{j+1}} \left\|\frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N} - \dot{\bar{x}}(t)\right\|\mathrm{d}t$$

$$=2\int_a^b \|\dot{x}_N(t)-\dot{\bar{x}}(t)\|\mathrm{d}t =: \nu_N \to 0 \quad (N\to\infty).$$

于是, 不失一般性, 可假设

$$\dot{x}_N(t)\to\dot{\bar{x}}(t), \quad \theta_N(t)\to 0, \quad \text{a.e. } t\in[a,b] \quad (N\to\infty).$$

沿着指定序列 $N\to\infty$ (不失一般性, 这里取整个自然数列), 考虑近似离散 Euler-Lagrange 包含 (6.41). 由 (6.41) 可找到

$$(x^*_{Nj}, v^*_{Nj}) \in \widehat{\partial}\vartheta_j\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N}\right), \quad j=0,\cdots,N-1$$

和 $e^*_{Nj}, \widetilde{e}^*_{Nj} \in \mathbb{B}^*$, 使得包含

$$\left(\frac{p_N(t_{j+1})-p_N(t_j)}{h_N} - \lambda_N x^*_{Nj}\right) + \varepsilon_N e^*_{Nj}$$

$$\in \widehat{D}^*F_j\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1})-\bar{x}_N(t_j)}{h_N}\right)\left(\lambda_N v^*_{Nj} + \lambda_N\frac{\theta_{Nj}}{h_N}b^*_{Nj} - p_N(t_{j+1}) + \varepsilon_N\widetilde{e}^*_{Nj}\right)$$

对所有 $j=0,\cdots,N-1$ 和所有 $N\in\mathbb{N}$ 都成立. 由 (H3′) 中 ϑ 的局部 Lipschitz 连续性和命题 1.85 可得到

$$\|(x^*_{Nj}, v^*_{Nj})\| \leqslant \ell_\vartheta, \quad \forall j=0,\cdots,N-1, \quad N\in\mathbb{N},$$

其中 ℓ_ϑ 是不依赖 $t\in[a,b]$ 的 $\vartheta(\cdot,\cdot,t)$ 的一致 Lipschitz 模. 由 (H1) 中 F 的 Lipschitz 连续性和定理 1.43 中的上导数条件, 可得到估计

$$\left\|\frac{p_N(t_{j+1})-p_N(t_j)}{h_N} - \lambda_N x^*_{Nj} + \varepsilon_N e^*_{Nj}\right\|$$

$$\leqslant \ell_F\left\|\lambda_N v^*_{Nj} + \lambda_N\frac{\theta_{Nj}}{h_N}b^*_{Nj} - p_N(t_{j+1}) + \varepsilon_N\widetilde{e}^*_{Nj}\right\|$$

对 $j=0,\cdots,N-1$ 成立. 类似问题 6.19 的证明且考虑到 $\|p_N(t_N)\|\leqslant 1$, 由这些估计可导出 $p_N(t)$ 在 $[a,b]$ 上一致有界且存在不依赖 N 的正数 α 和 β, 使得

$$\|\dot{p}_N(t)\| \leqslant \alpha + \beta\|\theta_N(t)\|, \quad \text{a.e. } t\in[a,b].$$

因为空间 X 和 X^* 都具有 RNP, 由 $L^1([a,b];X^*)$ 中关于弱紧性的 Dunford 定理, 可得到 $\{\dot{p}_N(\cdot)\}$ 的一个在该空间中弱收敛到某个 $v(\cdot)\in L^1([a,b];X^*)$ 的子列. 利用

Bochner 积分作为从 $L^1([a,b];X^*)$ 到 X^* 的线性算子的弱连续性和估计 $\|p_N(b)\| \leqslant 1$, 可断定存在一个绝对连续函数 $p\colon [a,b] \to X^*$ 满足

$$p(t) := p(b) + \int_t^b v(s)\mathrm{d}s, \quad a \leqslant t \leqslant b,$$

其中 $p(b)$ 是 $\{p_N(b)\}$ 在 X^* 中弱 * 拓扑下的一个极限点, 并且使得对所有 $t \in [a,b]$, $p_N(t)$ 在 X^* 中弱 * 收敛于 $p(t)$ (因此在该空间中是弱 * 收敛的). 于是, 经典的 Mazur 定理确保了当 $N \to \infty$ 时, $\{\dot{p}_N(\cdot)\}$ 的某凸组合序列在 $L^1([a,b];X^*)$ 中强收敛于 $\dot{p}(\cdot)$. 因此 $\dot{p}_N(\cdot)$ (的某子序列, 这里不再重新标识下标) 在 $[a,b]$ 上几乎处处收敛于 $\dot{p}(\cdot)$.

对任意给定的 $N \in \mathbb{N}$, 近似 Euler-Lagrange 包含 (6.41) 可改写为

$$\begin{aligned}\dot{p}_N(t) \in \Big\{u \in X^* \Big| (u, p_N(t_{j+1}) - \lambda_N \theta_N(t)) \in \lambda_N \widehat{\partial}\vartheta(\bar{x}_N(t_j), \dot{\bar{x}}_N(t), t_j)\\ + \widehat{N}((\bar{x}_N(t_j), \dot{\bar{x}}_N(t)); \operatorname{gph} F(t_j)) + \varepsilon_N \mathbb{B}^*\Big\}\end{aligned}$$

对 $t \in [t_j, t_{j+1})$ 和 $j = 0, \cdots, N-1$ 成立. 现在, 当 $N \to \infty$ 时, 取极限并利用上面建立的逐点收敛的结果, 即得广义 Euler-Lagrange 包含 (6.44).

为了导出横截性包含 (6.45), 当 $N \to \infty$ 时, 在离散横截性条件 (6.42) 中取极限. 唯一需要阐明的是, 从 $\Omega_N = \Omega + \eta_N \mathbb{B}$ 的 Fréchet法向量转化为 Ω 的基本法向量的可能性. 利用定理 3.7(i) 的和法则和当 $N \to \infty$ 时 $\eta_N \downarrow 0$ 的事实, 这容易办到.

剩下只需证明非平凡条件 $(\lambda, p(\cdot)) \neq 0$. 假设 $\lambda = 0$, 不失一般性, 对所有 $N \in \mathbb{N}$, 可令 $\lambda_N = 0$. 需要证明 $p(\cdot)$ 在 $[a,b]$ 上不恒为零. 如若不然, 即 $p(t) = 0$, $\forall\, t \in [a,b]$. 那么由上面的证明, 对所有 $t \in [a,b]$ 有 $p_N(t) \xrightarrow{w^*} 0$; 特别地, 当 $N \to \infty$ 时, $p_N(t_0) \xrightarrow{w^*} 0$ 和 $p_N(t_N) \xrightarrow{w^*} 0$. 在此情形, 离散横截性包含 (6.24) 可写为

$$(p_N(t_0), -p_N(t_N)) \in \widehat{N}((\bar{x}_N(t_0), \bar{x}_N(t_N)); \Omega + \eta_N \mathbb{B}) + \varepsilon_N \mathbb{B}^*. \tag{6.46}$$

对 (6.46) 中和的 Fréchet法锥再利用定理 3.7(i), 然后利用 Ω 在点 $(\bar{x}(a), \bar{x}(b))$ 关于第二个分量的强 PSNC 性质, 可得到当 $N \to \infty$ 时 $\|p_N(t_N)\| \to 0$, 这与定理 6.19 中的非平凡条件 (6.42) 相矛盾. 证毕. △

下面的定理给出了原 Bolza 问题 (P) 的必要最优性条件, 该条件是在离散时间问题 (P_N) 的近似必要最优性条件中取极限导出的. 与定理 6.21 相比, 该定理应用到可和被积函数 $\vartheta(x, v, \cdot)$ 上并且给出了较好的 Euler-Lagrange 包含形式. 另一方面, 它在所论状态空间 X 上施加了更多限制假设. 在该定理的阐述和证明中, 像前面讨论的定理 6.21 那样, 保持同样的记号约定.

定理 6.22(具有可和被积函数 Bolza 问题松弛局部极小点的广义 Euler-Lagrange 条件) 设 $\bar{x}(\cdot)$ 为 Bolza 问题 (P) 在假设 (H1), (H2), (H3$'$) 和 (H4$'$) 下的一

个松弛中间局部极小点. 假设空间 X 是自反和可分的且集合 Ω 在点 $(\bar{x}(a), \bar{x}(b))$ 关于第二个分量的强 PSNC 的. 那么存在 $\lambda \geqslant 0$ 和绝对连续映射 $p\colon [a,b] \to X^*$, 不都为零, 满足广义 Euler-Lagrange 包含

$$\begin{aligned}\dot{p}(t) \in \quad & \operatorname{co}\Big\{u \in X^* \Big| (u, p(t)) \in \lambda\partial\vartheta(\bar{x}(t), \dot{\bar{x}}(t), t) \\ & + N_+((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F(t))\Big\}, \quad \text{a.e. } t \in [a,b] \end{aligned} \tag{6.47}$$

和横截性包含 (6.45).

证明 沿定理 6.21 的证明思路, 利用离散逼近问题列 (P_N) 替换 $(\overline{P}_N)$. 唯一的区别在于广义 Euler-Lagrange 包含 (6.47) 的证明与 (6.44) 不同, 一般来说, 这两个包含分别基于不同的离散版本 (6.43) 和 (6.41), 证明中需要的假设也不太一样.

为此, 为了记号方便, 假设当 $N \to \infty$ 时, 无需在集值积分下取闭包, 离散 Euler-Lagrange 包含 (6.43) 即成立; 这在如下的证明中并不限制一般性. 那么由 (6.43) 和 Fréchet上导数的定义, 存在对偶元

$$(x_{Nj}^*, v_{Nj}^*) \in \int_{t_j}^{t_{j+1}} \partial\vartheta\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t\right), \quad j = 0, \cdots, N-1,$$

以及 $e_{Nj}^*, \widetilde{e}_{Nj}^* \in \mathbb{B}^*$ 满足包含

$$\begin{aligned}&\left(\frac{p_N(t_{j+1}) - p_N(t_j)}{h_N} - \lambda_N x_{Nj}^*\right) + \varepsilon_N e_{Nj}^* \\ &\in \widehat{D}^* F_j\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}\right)\left(\lambda_N v_{Nj}^* + \lambda_N \frac{\theta_{Nj}}{h_N} b_{Nj}^* - p_N(t_{j+1}) + \varepsilon_N \widetilde{e}_{Nj}^*\right)\end{aligned}$$

沿着 $N \to \infty$ 时的一个子序列对所有 $j = 0, \cdots, N-1$ 成立; 为简便起见, 以自然数序列代替该子序列. 由定理 6.21 的证明, 可找到一个绝对连续映射 $p\colon [a,b] \to X^*$, 使得在 X^* 中当 $N \to \infty$ 时, $p_N(t) \xrightarrow{w} p(t)$, $\forall t \in [a,b]$ 以及 $\dot{p}_N(t)$ 的一列凸组合在 $[a,b]$ 上几乎处处收敛于 $\dot{p}(t)$. 那么将上面的离散包含改写为

$$\begin{aligned}\dot{p}_N(t) \in \Big\{u \in X^* \Big| & (u, p_N(t_{j+1}) - \lambda_N\theta_N(t)) \in \frac{\lambda_N}{h_N}(x_{Nj}^*, v_{Nj}^*) \\ & + \widehat{N}((\bar{x}_N(t_j), \dot{\bar{x}}_N(t)); \operatorname{gph} F(t_j)) + \varepsilon_N \mathbb{B}^*\Big\}\end{aligned}$$

对 $t \in [t_j, t_{j+1})$ 和 $j = 0, \cdots, N-1$ 成立. 由 (x_{Nj}^*, v_{Nj}^*) 的构造存在可和映射 $u_{Nj}^*\colon [t_j, t_{j+1}] \to X^*$ 和 $w_{Nj}^*\colon [t_j, t_{j+1}] \to X^*$ 满足关系

$$(u_{Nj}^*(t), w_{Nj}^*(t)) \in \partial\vartheta\left(\bar{x}_N(t_j), \frac{\bar{x}_N(t_{j+1}) - \bar{x}_N(t_j)}{h_N}, t\right), \quad \text{a.e. } t \in [t_j, t_{j+1}],$$

$$\frac{(x_{Nj}^*, v_{Nj}^*)}{h_N} = \frac{1}{h_N}\int_{t_j}^{t_{j+1}} (u_{Nj}^*(t), w_{Nj}^*(t))\mathrm{d}t, \quad j = 0, \cdots, N-1.$$

在整个区间 $[a,b]$ 上, 定义映射序列 $u_N^*\colon [a,b]\to X^*$ 和 $w_N^*\colon [a,b]\to X^*$ 为

$$(u_N^*(t), w_N^*(t)) := (u_{Nj}^*(t), w_{Nj}^*(t)), \quad t\in[t_j,t_{j+1}), \quad j=0,\cdots,N-1.$$

因为 $u_N^*(\cdot)$ 和 $w_N^*(\cdot)$ 在 $[a,b]$ 上可积有界, 由 Dunford 定理, 存在子序列, 它们在 $L^1([a,b];X^*)$ 的弱拓扑下收敛于 $u^*(\cdot)$ 和 $w^*(\cdot)$, 再利用 Mazur 弱闭包定理以及定理 6.13 中的强收敛 $\bar{x}_N(\cdot)\to\bar{x}(\cdot)$, 有关系式

$$(u^*(t), w^*(t)) \in \operatorname{cl}\operatorname{co}\partial\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t) = \operatorname{co}\partial\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t), \quad \text{a.e. } t\in[a,b],$$

其中由于 X 的自反性和 $\operatorname{co}\partial\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t)$ 在 X^* 中弱拓扑下的紧性, 于是该集合在 X 的强拓扑下是闭的, 故闭包运算可以略去. 现在利用引理 6.18 证明提到的无穷维 Lyapunov-Aumann 定理, 以及众所周知的关于 Bochner 积分的性质

$$\lim_{h\to 0}\frac{1}{h}\int_t^{t+h} f(s)\mathrm{d}s = f(t) \quad (\text{a.e. } t\in[a,b])$$

和自反空间中局部 Lipschitz 函数的基本次微分的弱闭性 (比照定理 3.59), 可断言存在 $\vartheta(\cdot,\cdot,t)$ 的次梯度 $(x^*(t), v^*(t))$ 使得

$$\frac{\lambda_N}{h_N}(x_{Nj}^*, v_{Nj}^*) \xrightarrow{w^*} (x^*(t), v^*(t)) \in \partial\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t), \quad \text{a.e. } t\in[a,b].$$

最后当 $N\to\infty$ 时, 在上面关于 $\dot{p}_N(\cdot)$ 的包含中取极限, 即可得到所要的广义 Euler-Lagrange 包含 (6.47), 其中由于 $p_N(\cdot)$ 和 $\dot{p}_N(\cdot)$ 的一致有界性, 在所考虑的自反情形, 闭包运算可以去掉; 参见上面的讨论. 注意到, 在所论情形只需用到 (6.47) 中被积函数 $\vartheta(\cdot,\cdot,t)$ 的基本次微分, 而不必用 (6.44) 中的增广次微分. 证毕. △

定理 6.21 和定理 6.22 的非平凡条件确保了满足 Euler-Lagrange 和横截性包含的 $(\lambda, p(\cdot))$ 不为零. 下面的结果在额外的假设下给出了加强的非平凡条件, 即 $(\lambda, p(b))\neq 0$.

推论 6.23(具有加强非平凡性的广义 Euler-Lagrange 条件) 设 $\bar{x}(\cdot)$ 为 Bolza 问题 (P) 的松弛中间局部极小点. 除了定理 6.21 和定理 6.22 的假设外, 假设

(a) $\Omega=\Omega_a\times\Omega_b$, 其中 Ω_b 在点 $\bar{x}(b)$ 是 SNC 的; 或者

(b) Ω 在点 $(\bar{x}(a),\bar{x}(b))$ 关于第二个分量是强 PSNC 的, $F(\cdot,t)$ 在点 $(\bar{x}(t),\dot{\bar{x}}(t))$ 是强上导数正规的并且对几乎处处 $t\in[a,b]$, $\operatorname{gph} F(t)$ 在该点是法向半连续的.

那么有广义 Euler-Lagrange 和横截性包含 (6.44) 和 (6.45) (或 (6.47) 和 (6.45)), 并且具有加强的非平凡条件 $\lambda+\|p(b)\|=1$, 其中对情形 (b), 用 $N((\bar{x}(t),\dot{\bar{x}}(t));\operatorname{gph} F(t))$ 替换 $N_+((\bar{x}(t),\dot{\bar{x}}(t));\operatorname{gph} F(t))$.

证明 由定理 6.21 和定理 6.22 中非平凡条件 (相同) 的证明, 在离散近似中可得伴随轨道 $p_N(\cdot)$ 的横截性包含 (6.46) 且 $\lambda_N=0$. 假设 (a) 成立. 那么可得

$$-p_N(t_N) \in \widehat{N}(\bar{x}_N(t_N);\Omega_b+\eta\mathbb{B}) + \varepsilon_N\mathbb{B}^* \quad (N\to\infty),$$

由定理 3.7(i) 和 Ω_b 在点 $\bar{x}(b)$ 的 SNC 性质, 这意味着只要 $p_N(t_N) \xrightarrow{w^*} 0$ (当 $N \to \infty$ 时), 就有 $\|p_N(t_N)\| \to 0$. 分别由定理 6.19 和定理 6.20, 这显然和离散时间问题 $(\overline{P}_N)$ 和 (P_N) 的非平凡条件相矛盾.

剩下来, 在情形 (b) 中证明非平凡条件 $\lambda + \|p(b)\| \neq 0$. 这可由如下的事实得到: 只要 $p(\cdot)$ 满足广义 Euler-Lagrange 包含 (6.44) 并且 $\lambda = 0$ 和 $p(b) = 0$, 那么 $p(t) = 0$ 对所有 $t \in [a,b]$ 成立. 事实上, 在此情形中, 利用 $\mathrm{gph}\, F(t)$ 的法向半连续性, 可将 (6.44) 写为

$$\dot{p}(t) \in \mathrm{cl\,co}\left\{u \in X^* \middle| (u, p(t)) \in N((\bar{x}(t), \dot{\bar{x}}(t)); \mathrm{gph}\, F(t))\right\}, \quad \text{a.e. } t \in [a,b].$$

由 (b) 中的强上导数正规性假设, 它等价于

$$\dot{p}(t) \in \mathrm{cl\,co}\, D_M^* F(\bar{x}(t), \dot{\bar{x}}(t))(-p(t)), \quad \text{a.e. } t \in [a,b].$$

由定理 1.44 中 Lipschitz 连续性的混合上导数条件, 后者明显地蕴涵: 当 $p(b) = 0$ 时, 在 $[a,b]$ 上有 $p(t) \equiv 0$. 证毕. △

如果 X 是有限维的, 那么在每一点, 任何集合是 SNC 的且任何映射 $F\colon X \rightrightarrows X$ 是强上导数正规的. 因此自动地有定理 6.22 和推论 6.23 中的广义 Euler-Lagrange 条件. 在约束集 Ω 上, 另外一种不需要任何 SNC/PSNC 假设的情形是端点约束由涉及局部 Lipschitz 函数的有限多个等式和不等式给出的情形, 该情形将在下面考虑.

推论 6.24(具有泛函端点约束问题的广义 Euler-Lagrange 条件) 设问题 (P) 中的端点约束 Ω 由

$$\begin{aligned} \Omega := \Big\{(x_a, x_b) \in X^2 \Big|\ & \varphi_i(x_a, x_b) \leqslant 0, \quad i = 1, \cdots, m, \\ & \varphi_i(x_a, x_b) = 0, \quad i = m+1, \cdots, m+r\Big\} \end{aligned}$$

给出, 其中每一个 φ_i 和费用函数 $\varphi_0 := \varphi$ 在 $(\bar{x}(a), \bar{x}(b))$ 附近是局部 Lipschitz 的. 假设推论 6.23 中除了那些涉及 Ω 的 SNC/PSNC 性质之外的所有条件都成立. 那么存在非负乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$ 使得

$$\lambda_i \varphi_i(\bar{x}(a), \bar{x}(b)) = 0, \quad i = 1, \cdots, m,$$

以及存在绝对连续伴随轨线 $p\colon [a,b] \to X^*$ 满足推论 6.23 中的广义 Euler-Lagrange 包含和下面的横截性条件:

$$\begin{aligned} (p(a), -p(b)) \in & \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}(a), \bar{x}(b)) \\ & + \sum_{i=m+1}^{m+r} \lambda_i \Big[\partial\varphi_i(\bar{x}(a), \bar{x}(b)) \bigcup \partial(-\varphi_i)(\bar{x}(a), \bar{x}(b))\Big]. \end{aligned}$$

特别地, 如果所有的 φ_i 在点 $(\bar{x}(a),\bar{x}(b))$ 是严格可微的, 那么存在 $(\lambda_0,\cdots,\lambda_{m+r})\neq 0$ 满足上面的互补松弛条件和标准的符号条件

$$\lambda_i \geqslant 0 \quad (i=0,\cdots,m),$$

并且使得推论 6.23 中相应的 Euler-Lagrange 包含以及横截性条件

$$(p(a),-p(b))=\sum_{i=m+1}^{m+r}\lambda_i\nabla\varphi_i(\bar{x}(a),\bar{x}(b))$$

成立.

证明 首先假设局部 Lipschitz 函数 $\varphi_1,\cdots,\varphi_{m+r}$ 满足定理 3.86 中非光滑的 Mangasarian-Fromovitz 约束规范条件. 那么该推论中定义的约束集合 Ω 在点 $(\bar{x}(a),\bar{x}(b))$ 是 SNC 的. 更进一步, 对具体的 $F:=(\varphi_1,\cdots,\varphi_{m+r})$ 和

$$\Theta:=\Big\{(\alpha_1,\cdots,\alpha_{m+r})\in\mathbb{R}^{m+r}\Big|\alpha_i\leqslant 0,\ i=1,\cdots,m,\\ \alpha_i=0,\ i=m+1,\cdots,m+r\Big\}$$

由定理 3.8 的分析法则可得相同的约束规范条件, 从而确保包含

$$N(\bar{z};\Omega)\subset\Big\{\sum_{i=1}^{m}\lambda_i\partial\varphi_i(\bar{z})+\sum_{i=m+1}^{m+r}\lambda_i\left[\partial\varphi_i(\bar{z})\cup\partial(-\varphi_i)(\bar{z})\right]\Big|\\ \lambda_i\geqslant 0,\ i=1,\cdots,m+r;\quad \lambda_i\varphi_i(\bar{z})=0,\ i=1,\cdots,m\Big\}$$

对约束集合 Ω 在点 $\bar{z}:=(\bar{x}(a),\bar{x}(b))$ 的基本法向量成立. 那么由 $\lambda_0=\lambda$ 时的 (6.45) 可得该推论的横截性包含, 其中的非平凡条件 $(\lambda,p(b))\neq 0$ 等价于 $(\lambda_0,\cdots,\lambda_{m+r})\neq 0$. 最后假设定理 3.86 中的规范条件不成立, 立即可得到所要的横截性包含并且 $(\lambda_1,\cdots,\lambda_{m+r})\neq 0$. 证毕. △

注意到, 如果约束集合 Ω 不是 SNC (或强 PSNC) 的, 特别地, 当定理 3.86 中 Mangasarian-Fromovitz 型约束规范条件不满足时, 由推论 6.23 启发的加强的非平凡条件 $(\lambda_0,p(b))\neq 0$ 在推论 6.24 的框架下可能不成立. 例如, 对某些具有 $x(a)=x_0$ 和 $x(b)=x_1$ 光滑抛物系统的最优控制的两点边界问题; 参见 Fattorini [432] 及 Li 与 Yong [789] 中众所周知的例子. 另一方面, 当 $x(a)=x_0$ 和 $x(b)\in x_1+r\mathbb{B}$ 且 $r>0$ 时, 推论 6.23(a) 中的 SNC 要求得到满足, 因为后者作为球总是 SNC 的 (由命题 1.25, 它实际上是上图 Lipschitz 的).

也注意到, 利用 Fréchet次梯度的光滑变分描述 (类似定理 5.19 对不可微规划问题的证明) 和利用对光滑端点函数情形的推论 6.24 的结果, 可得到相应于定理

6.21 和定理 6.22 的具有上次微分横截性条件的结果; 参见注 6.30 中的确切阐述和更多的细节.

作为本节的结束, 讨论一些特别的问题, 大多和上面具有无穷维状态空间微分包含的 Euler-Lagrange 条件相关.

注 6.25(关于 Euler-Lagrange 条件的讨论)

(i) 由定理 6.21 和 6.22 的证明可知, 为了确保非平凡条件, Ω 上的强 PSNC 假设可替换为下面关于 F 的另外一个假设: 存在 $t \in [a,b]$, 使得对任何序列 $t_k \to t$, $x_k \to \bar{x}(t)$, $v_k \in F(x_k,t_k)$ 和 $(x_k^*,v_k^*) \in \widehat{N}((x_k,v_k);\mathrm{gph}\,F(t_k))$, 有

$$(x_k^*,v_k^*) \xrightarrow{w^*} (0,0) \Rightarrow \|v_k^*\| \to 0 \quad (k \to \infty).$$

该性质和 F 在点 $(\bar{x}(t),t)$ 关于像分量的强 PSNC 性质密切关联; 也可参见定义 5.71 中移动集合的关于 SNC 的类似概念.

(ii) 由定理 1.21, 具有非空相对内点集的凸集的 SNC 性质等价于其闭仿射包的有限余维数性质. 强 PSNC 性质本质上要比 SNC 弱; 参见定理 1.75.

(iii) 如果速率集合 $F(x,t)$ 和被积函数 $\vartheta(x,\cdot,t)$ 在给定的局部极小点附近是凸的, 那么定理 6.21 中 Euler-Lagrange 包含明显蕴涵 Weierstrass-Pontryagin 最大值条件

$$\langle p(t),\dot{\bar{x}}(t)\rangle - \lambda\vartheta(\bar{x}(t),\dot{\bar{x}}(t),t) = \max_{v \in F(\bar{x}(t),t)} \big\{ \langle p(t),v\rangle - \lambda\vartheta(\bar{x}(t),v,t)\big\}$$

对几乎处处 $t \in [a,b]$ 成立. 它可由定理 1.34 中凸集值映射上导数的极值性质直接导出. 在最优控制中, 后者是“Pontryagin 最大值原理”框架下的主要条件. 在下一小节中将证明, 至少对自反和可分状态空间在一些额外附加的条件下, 最大值条件是广义 Euler-Lagrange 包含没有凸性要求时的补充. 沿此思路, 注意到定理 6.21 和定理 6.22 中要求的 SNC (实际上是强 PSNC) 性质可视为有限余维数条件的非凸变体. 这个余维数条件出现在 (6.2) 类型发展方程控制问题及其在 PDE 上的具体化 (具有光滑速率映射 f 和凸约束/目标集 Ω) 的必要最优条件理论中; 参见前面提到的书 Fattorini [432] 及 Li 与 Yong [789] 以及其中的参考文献和讨论.

注 6.26(半线性无界微分包含的最优控制) 许多涉及半线性偏微分方程的重要模型可由 $\mathcal{C}_0$ 半群来描述; 这里再次提到 Fattorini [432] 及 Li 与 Yong [789] 以及本书 7.2~7.4 节中的材料. 这样一来, 本节中的控制问题 (P) 的一个类似问题是用发展型模型 $\dot{x}(t) \in Ax(t) + F(x(t),t)$ 替换微分包含 (6.1), 其中 A 是 X 上的紧 $\mathcal{C}_0$ 半群的无穷小生成元且该包含的连续解 $x(\cdot)$ 是在适度解意义下的. 适度解的意思是说存在一个 Bochner 可积映射 $v(\cdot) \in L^1([a,b];X)$, 使得

$$v(t) \in F(x(t),t), \quad \text{a.e. } t \in [a,b],$$

$$x(t) = \mathrm{e}^{A(t-a)}x(a) + \int_a^t \mathrm{e}^{A(t-s)}v(s)\mathrm{d}s, \quad t \in [a,b].$$

对于 Mayer 费用函数情形

$$\begin{aligned} &\min && \varphi(x(a), x(b)) \\ &\text{s.t.} && (x(a), x(b)) \in \Omega \subset X^2, \end{aligned}$$

由上面发展的方法, 在最优解附近对速率集 $F(x,t)$ 额外附加凸性的假设下, 即可导出必要最优性条件. 那么对自反和可分的状态空间 X 和自治系统 (为简单起见) 的情形, 广义 Euler-Lagrange 包含可阐述如下:

$$\left\{ \begin{aligned} p(t) \in &\,\mathrm{e}^{A^*(b-t)}p(b) \\ &+ \int_b^t \left\{ \mathrm{e}^{A^*(s-t)} D_N^* F(\bar{x}(s), v)(-p(s)) \Big| v \in M(\bar{x}(s), p(s)) \right\} \mathrm{d}s \end{aligned} \right.$$

对所有 $t \in [a,b]$ 成立, 其中 p: $[a,b] \to X^*$ 是连续映射, 它满足横截性和非平凡条件

$$(p(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega), \quad \lambda + \|p(b)\| \neq 0,$$

并且 $\lambda \geqslant 0$, 其中极大点集 $M(x,p)$ 定义为

$$M(x,p) := \{v \in F(x) \mid \langle p, v\rangle = \mathcal{H}(x,p)\}$$

并且

$$\mathcal{H}(x,p) := \max\{\langle p, v\rangle \mid v \in F(x)\}.$$

此外, 在此情形广义 Euler-Lagrange 包含蕴涵 Weierstrass-Pontryagin 最大值条件

$$\langle p(t), \bar{v}(t)\rangle = \mathcal{H}(\bar{x}(t), p(t)), \quad \text{a.e. } t \in [a,b]$$

其中可测映射 $\bar{v}(t) \in F(\bar{x}(t))$ 满足

$$p(t) \in \mathrm{e}^{A^*(b-t)}p(b) + \int_b^t \left\{ \mathrm{e}^{A^*(s-t)} D_N^* F(\bar{x}(s), \bar{v}(s))(-p(s)) \right\} \mathrm{d}s, \quad t \in [a,b];$$

参见 Mordukhovich 与 D. Wang[970, 971] 文章中的证明和相关结果的更多讨论.

6.2 无松弛微分包含的必要最优性条件

本节主要专注于导出无任何松弛非凸微分包含的必要最优性条件, 它是基于由一族无微分包含和无端点约束的 Bolza 问题来逼近原约束问题. 与前一节建立的一

般结果相比, 无约束 Bolza 问题的广义 Euler-Lagrange 条件和所作的假设允许实质上的具体化. 通过取极限, 可以得到具有自反和可分状态空间的原控制问题的任意 (即非松弛) 中间极小点的 Euler-Lagrange 类型的必要最优性条件. 此外, 它们是一般非凸情形中 Weierstrass-Pontryagin 最大值条件的补充. 如果状态空间 X 是有限维的且速率集合 $F(x,t)$ 是凸的, 那么上面的 Euler-Lagrange 和最大值条件等价于增广 Hamilton 包含, 该包含是通过与 $F(x,t)$ 相关联的 Hamilton 函数的基本次微分的偏凸化来表示的. 本节也讨论所得结果的各种推广并给出一些解说性的例子.

6.2.1 中间局部极小点的 Euler-Lagrange 和最大值条件

与定理 6.22 相比, 为实现上面提到的方法, 需要对初始数据作一些额外的假设, 但速率映射 $F(x,\cdot)$ 的几乎处处连续性假设可替换为可测性假设. 更进一步, 本节考虑前一节中研究过的问题 (P) 的如下 Mayer 形式 (P_M), 它具有左端点固定的可行轨线: 在微分包含

$$\dot{x}(t) \in F(x(t),t), \quad \text{a.e. } t \in [a,b], \quad x(a) = x_0 \tag{6.48}$$

的绝对连续轨道上

$$\min \quad \varphi(x(b)) \quad \text{s.t.} \quad x(b) \in \Omega \subset X.$$

通过标准状态增广技巧, Bolza 问题中非零被积函数 f 的一般情形可化为 Mayer 问题. 也注意到, 由于下面讨论中的状态空间 X 假设为自反和可分的, (6.48) 中绝对连续解的概念和定义 6.1 中的一致.

首先描述 (6.48) 中关于集值映射 F 的假设. 这比定理 6.22 中要弱. 保持 6.1.1 小节中紧性和 F 关于 x 的 Lipschitz 连续性的假设 (H1), 在 $[a,b]$ 上有时可能需要具有可和函数 $m_F(\cdot)$ 和 $\ell_F(\cdot)$ (在某些方向上, 由各种标准的约化, 这些条件也可以放宽, 例如见文献 [255, 261, 598, 1289]), 将 F 关于时间变量 $t \in [a,b]$ 的几乎处处连续性假设 (H2) 替换为可测性假设. 注意到, 可分空间中具有闭值集值映射的所有合理的可测性概念都是等价的 (参见引理 6.18 证明中的讨论), 在本节中亦如此.

(H2′) $F(x,\cdot)$ 在区间 $[a,b]$ 上是可测的, 且关于 x 在 (H1) 中的开集 $U \subset X$ 上是一致的.

另外, 费用函数 $\varphi = \varphi(x)$ 在 (H4) 和 (H4′) 中的连续性和 Lipschitz 连续性假设也可减弱, 因为减弱的假设就足以得出 Mayer 问题改进 (更一般) 的横截性条件了. 也就是说, 将 (H4) 和 (H4′) 替换为下面的假设:

(H4″) 相对于 Ω, φ 在 $\bar{x}(b)$ 附近是下半连续的, 而 Ω 在该点附近是局部闭的.

另一方面, 与定理 6.22 和推论 6.23 相比, 下面的定理在 F 上附加了上导数正规性和 SNC 的假设. 注意到, 下面给出的广义 Euler-Lagrange 包含的上导数形式等价

于推论 6.23 在 $\vartheta=0$ 的情形, 但这里不需对 $\operatorname{gph}F(t)$ 作法向半连续假设. 本小节的其余部分研究定义 6.7 中秩为 1 的中间局部极小点. 依照惯例 $\varphi_\Omega(\cdot)=\varphi(\cdot)+\delta(\cdot;\Omega)$.

定理 6.27(非凸微分包含的 Euler-Lagrange 和 Weierstrass-Pontryagin 条件) 设 $\bar{x}(\cdot)$ 为 Mayer 问题 (P_M) 在假设 (H1), (H2′) 和 (H4″) 下的中间局部极小点. 另外假设:

(a) Banach 空间 X 是自反的、可分的且具有等价的 Kadec 重赋范;

(b) 函数 φ_Ω 在点 $\bar{x}(b)$ 是 SNEC 的且它的上图是弱闭的;

(c) 映射 $F(\cdot,t)\colon X\rightrightarrows X$ 在点 $(\bar{x}(t),\dot{\bar{x}}(t))$ 是 SNC 的, 在该点附近是强上导数正规的且它的图像对几乎处处 $t\in[a,b]$ 是弱闭的.

那么存在数 $\lambda\geqslant 0$ 和绝对连续伴随轨线 $p\colon[a,b]\to X^*$, 不都为零, 满足 Euler-Lagrange 包含

$$\dot{p}(t)\in\operatorname{co}D_x^*F(\bar{x}(t),\dot{\bar{x}}(t),t)(-p(t)),\quad \text{a.e. } t\in[a,b], \tag{6.49}$$

Weierstrass-Pontryagin 最大值条件

$$\langle p(t),\dot{\bar{x}}(t)\rangle=\max_{v\in F(\bar{x}(t),t)}\langle p(t),v\rangle,\quad \text{a.e. } t\in[a,b], \tag{6.50}$$

和横截性包含

$$(-p(b),-\lambda)\in N((\bar{x}(b),\bar{\beta});\operatorname{epi}\varphi_\Omega). \tag{6.51}$$

此外, (6.51) 总是蕴涵

$$-p(b)\in\partial[\lambda\varphi+\delta(\cdot;\Omega)](\bar{x}(b)). \tag{6.52}$$

如果 φ 在 $\bar{x}(b)$ 附近相对于 Ω 是 Lipschitz 连续的, 那么 (6.51) 和 (6.52) 等价.

证明 在无其他约束的原微分包含 (6.1) 的可行轨线/轨道上考虑参数泛函

$$\theta_\beta(x):=\operatorname{dist}((x(b),\beta);\operatorname{epi}\varphi_\Omega),\quad \beta\in\mathbb{R}.$$

下面关于 $\bar{x}(\cdot)$ 固定 (H1) 中的开集 $U\subset X$. 对每个 $\beta\in\mathbb{R}$, 显然有

$$\theta_\beta(\bar{x})\leqslant|\beta-\bar{\beta}|,$$

其中 β 充分靠近 $\bar{\beta}=\varphi(\bar{x}(b))$. 因为 $\bar{x}(\cdot)$ 是 (P_M) 的一个中间局部极小点并且由 $\theta_\beta(x)$ 的结构, 可得

$$\theta_\beta(x)>0,\quad\forall\,\beta<\bar{\beta},$$

只要 (6.48) 的轨道 $x(t)$ 属于局部极小点的某 $W^{1,1}$- 邻域且使得

$$x(t)\in U,\quad\forall\,t\in[a,b].$$

现在把 (6.48) 的满足唯一约束 $x(t) \in \mathrm{cl}\, U$ $(\forall\ t \in [a,b])$ 的所有轨道 $x(\cdot)$ 组成空间 $\mathcal{X}$, 并定义度量

$$d(x,y) := \int_a^b \|\dot{x}(t) - \dot{y}(t)\| \mathrm{d}t.$$

由原微分包含解的定义 6.1 和 Bochner 积分的标准性质易见, 度量空间 $\mathcal{X}$ 是完备的且对任何 $\beta \in \mathbb{R}$, 函数 $\theta_\beta(\cdot)$ 在 $\mathcal{X}$ 上是 (Lipschitz) 连续的. 由上面的构造可知, 对每一个 $\varepsilon > 0$, 存在 $\beta_\varepsilon < \bar{\beta}$, 使得 $\beta_\varepsilon \to \bar{\beta}$ (当 $\varepsilon \downarrow 0$ 时) 且

$$0 \leqslant \theta_\varepsilon(\bar{x}) < \varepsilon \leqslant \inf_{x \in \mathcal{X}} \theta_\varepsilon(x) + \varepsilon, \quad \text{其中 } \theta_\varepsilon := \theta_{\beta_\varepsilon}.$$

利用定理 2.26(i) 中的 Ekeland 变分原理, 可找到轨线 $x_\varepsilon \in \mathcal{X}$ 满足

$$d(x_\varepsilon, \bar{x}) \leqslant \sqrt{\varepsilon}, \quad \theta_\varepsilon(x) + \sqrt{\varepsilon} d(x, x_\varepsilon) \geqslant \theta_\varepsilon(x_\varepsilon)$$

对所有 $x \in \mathcal{X}$ 成立. 注意到, 上面的距离估计给出了 $x_\varepsilon(t) \in U$ $(\forall\ t \in [a,b])$ 且对小的 $\varepsilon > 0$, $x_\varepsilon(\cdot)$ 属于中间局部极小点 $\bar{x}(\cdot)$ 的固定 $W^{1,1}$- 邻域. 因此 $\theta_\varepsilon(x_\varepsilon) > 0$.

其次, 给定任意的 $\alpha, \varepsilon > 0$ 和 (6.5) 中的可和 Lipschitz 常数 $\ell_F(\cdot)$, 在所有满足 $x(t) \in U$ $(\forall\ t \in [a,b])$ 的绝对连续映射 $x\colon [a,b] \to X$ (无需是 (6.48) 的轨道) 的集合上, 定义 Bolza 型泛函

$$J_\varepsilon^\alpha[x] := \theta_\varepsilon(x) + \sqrt{\varepsilon} d(x, x_\varepsilon) + \alpha \int_a^b \sqrt{1 + \ell_F^2(t)}\, \mathrm{dist}\Big(\big(x(t), \dot{x}(t)\big); \mathrm{gph}\, F(t)\Big) \mathrm{d}t,$$

为此, 需要如下的辅助结果.

断言 存在常数 $\alpha \geqslant 1$ 使得对每个 $\varepsilon \in (0, 1/\alpha)$, 上面构造的绝对连续映射 $x_\varepsilon\colon [a,b] \to X$ 是 Bolza 泛函 J_ε^α 在约束

$$x(a) = x_0 \quad \text{和} \quad x(t) \in U \quad (\forall\ t \in [a,b])$$

下的一个中间局部极小点.

为了证明该论断, 首先注意到存在正数 ν, γ, 使得对每一个满足 $y(a) = x_0$, $y(t) \in U$ $(\forall\ t \in [a,b])$ 及

$$\int_a^b \mathrm{dist}\Big(\dot{y}(t); F\big(y(t), t\big)\Big) \mathrm{d}t < \nu$$

的轨线 $y(\cdot)$, 存在 (6.28) 的一个轨道 $x(\cdot)$ 满足

$$d(x,y) \leqslant \gamma \int_a^b \sqrt{1 + \ell_F^2(t)}\, \mathrm{dist}\Big(\big(y(t), \dot{y}(t)\big); \mathrm{gph}\, F(t)\Big) \mathrm{d}t. \tag{6.53}$$

事实上, 由关于微分包含拟轨道的 Filippov 定理 (参见 Aubin 与 Cellina [50] 第 120 页中的定理 1, 在 (H1) 和 (H2′) 中的假设下, 其证明对于无穷维包含也成立) 和估计

$$\operatorname{dist}(v, F(u,t)) \leqslant \sqrt{1+\ell_F^2(t)}\operatorname{dist}\Big((u,v);\operatorname{gph}F(t)\Big)$$

(在 (H1) 的假设下显然成立), 直接可得 (6.53). 现在假设上述论断不成立. 那么对每个 $k \in \mathbb{N}$, 存在 $\varepsilon_k \in (0,1/k)$ 和轨线 $y_k(\cdot) \in \mathcal{X}$ 满足 $y_k(t) \in U$, $\forall\, t \in [a,b]$,

$$\max_{t\in[a,b]}\|y_k(t) - x_{\varepsilon_k}(t)\| + \int_a^b \|\dot{y}_k(t) - \dot{x}_{\varepsilon_k}(t)\|\mathrm{d}t < \frac{1}{k},$$

以及 $J^k_{\varepsilon_k}[x_{\varepsilon_k}] > J^k_{\varepsilon_k}[y_k]$. 因此, 在 $W^{1,1}([a,b];X)$ 中的范数拓扑下 $y_k(\cdot) \to \bar{x}(\cdot)$. 此外,

$$J^k_{\varepsilon_k}[x_{\varepsilon_k}] = \theta_{\varepsilon_k}(x_{\varepsilon_k}) \downarrow 0 \quad (k \to \infty).$$

因此, 给定任何 $\nu > 0$, 对大的 k 有

$$\int_a^b \operatorname{dist}\Big(\dot{y}_k(t); F(y_k(t),t)\Big)\mathrm{d}t < J^k_{\varepsilon_k}[x_{\varepsilon_k}] < \nu.$$

由 (6.53), 这蕴涵着存在不依赖 k 的数 $\gamma > 0$ 和 (6.28) 的轨道 $x_k(\cdot)$ 使得当 $k \to \infty$ 时有

$$d(x_k, y_k) \leqslant \gamma \int_a^b \sqrt{1+\ell_F^2(t)}\operatorname{dist}\Big((y_k(t),\dot{y}_k(t));\operatorname{gph}F(t)\Big)\mathrm{d}t. \tag{6.54}$$

因为 (6.54) 的右端收敛于零并且在 $W^{1,1}([a,b];X)$ 中 $y_k(\cdot)$ 强收敛到 $\bar{x}(\cdot)$, 所以 $x_k(\cdot)$ 亦在 $W^{1,1}([a,b];X)$ 中强收敛到 $\bar{x}(\cdot)$, 这确保了对大的 $k \in \mathbb{N}$, 所有轨道 $x_k(\cdot) \in \mathcal{X}$ 属于中间局部极小点 $\bar{x}(\cdot)$ 的固定 $W^{1,1}([a,b];X)$- 邻域. 于是有

$$\begin{aligned} J^k_{\varepsilon_k}[x_k] \geqslant J^k_{\varepsilon_k}[x_{\varepsilon_k}] > J^k_{\varepsilon_k}[y_k] &= \theta_{\varepsilon_k}(y_k) + \sqrt{\varepsilon_k}d(y_k, x_{\varepsilon_k}) \\ &+ k\int_a^b \operatorname{dist}\Big(\dot{y}_k(t); F(y_k(t),t)\Big)\mathrm{d}t =: k\xi_k. \end{aligned}$$

现在考虑到 (6.54) 和 θ_ε 的构造, 对大的 k 有

$$k\xi_k < \sqrt{\varepsilon_k}\Big(d(x_k, x_{\varepsilon_k}) - d(y_k, x_{\varepsilon_k})\Big) + \theta_{\varepsilon_k}(x_k) - \theta_{\varepsilon_k}(y_k) \leqslant 3\gamma\xi_k,$$

矛盾. 这就完成了论断的证明.

注意到, 由于 U 在 X 中是开的, 从必要最优条件的角度来看, 约束 $x(t) \in U$ ($\forall\, t \in [a,b]$) 可忽略. 因此, 可将 $x_\varepsilon(\cdot)$ 视为下述具有有限值和 Lipschitz 数据的无约束 Bolza 问题的中间局部极小点:

$$\min \quad \varphi_\varepsilon(x(b)) + \int_a^b \vartheta_\varepsilon(x(t),\dot{x}(t),t)\mathrm{d}t, \tag{6.55}$$

其中要求绝对连续轨线 $x\colon [a,b] \to X$ 满足 $x(a) = x_0$, 并且在 $\bar{x}(\cdot)$ 的 $W^{1,1}$- 邻域内, 而端点费用函数由

$$\varphi_\varepsilon(x) := \operatorname{dist}((x, \beta_\varepsilon); \operatorname{epi} \varphi_\Omega) \tag{6.56}$$

给出, 并且被积函数为

$$\vartheta_\varepsilon(x, v, t) := \alpha\sqrt{1 + \ell_F^2(t)} \operatorname{dist}\Big((x, v); \operatorname{gph} F(t)\Big) + \sqrt{\varepsilon}\|v - \dot{x}_\varepsilon(t)\|. \tag{6.57}$$

注意到, 无约束问题 (6.55) 的任何中间局部极小点都是该问题的一个松弛中间局部极小点. 这可从定理 6.11 的松弛结果及其针对中间极小点的修正版本得到, 该修正由 Ioffe 与 Rockafellar[616] 的定理 4 给出, 它在无穷维空间中在所作的假设下是成立的. 也注意到假设 (H1), (H2′) 和 (H3″) 确保具有数据 (6.56) 和 (6.57) 的问题 (6.55) 满足定理 6.22 中除了速率集紧性外的所有假设条件. 事实上, 该紧性条件在 (6.55) 的无约束和 $W^{1,1}$- 有界的框架下是不需要的; 参见定理 6.22 的证明和前面的结果.

现在对任何固定的 $\varepsilon > 0$, 将定理 6.22 中的必要最优性条件用到问题 (6.55) 中. 首先利用具有 (6.57) 中被积函数 ϑ_ε 的广义 Euler-Lagrange 包含 (6.47), 然后利用定理 2.33(c) 中的和法则, 找到一个绝对连续伴随轨线 $p_\varepsilon\colon [a,b] \to X^*$ 满足

$$\begin{aligned} \dot{p}_\varepsilon(t) \in \operatorname{co}\Big\{u \in X^* \Big| (u, p_\varepsilon(t)) \in \mu(t)\partial \operatorname{dist}\Big((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)\Big) \\ + \sqrt{\varepsilon}(0, \mathbb{B}^*)\Big\} \quad \text{a.e. } t \in [a,b], \end{aligned}$$

其中 $\mu(t) := \alpha\sqrt{1 + \ell_F^2(t)}$. 固定 $t \in [a,b]$, 考虑关于 $(x_\varepsilon(t), \dot{x}_\varepsilon(t))$ 的两种情形:

(i) $\dot{x}_\varepsilon(t) \in F(x_\varepsilon(t), t)$ 和

(ii) $\dot{x}_\varepsilon(t) \notin F(x_\varepsilon(t), t)$.

在情形 (i) 中, 利用关于距离函数在集合内的点的次梯度的定理 1.97 得到近似伴随包含

$$\dot{p}_\varepsilon(t) \in \operatorname{co}\Big\{u \in X^* \Big| (u, p_\varepsilon(t)) \in N\Big((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)\Big) + \sqrt{\varepsilon}(0, \mathbb{B}^*)\Big\}.$$

考虑情形 (ii), 在 (a) 中关于 X 的 Kadec 范数结构假设下, 对距离函数在集合外的点的基本次梯度利用定理 1.105 中的第一个投影公式 (参见推论 1.106), 有包含

$$\partial \operatorname{dist}\Big((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)\Big) \subset \bigcup_{(x,v) \in \Pi\left((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)\right)} N\Big((x, v); \operatorname{gph} F(t)\Big).$$

考虑到逐点收敛 $(x_\varepsilon(t), \dot{x}_\varepsilon(t)) \to (\bar{x}(t), \dot{\bar{x}}(t))$ (当 $\varepsilon \downarrow 0$ 时), 有

$$\partial \operatorname{dist}\Big((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)\Big) \subset N\Big((\widetilde{x}_\varepsilon, \widetilde{v}_\varepsilon); \operatorname{gph} F(t)\Big)$$

对某收敛于 $(\bar{x}(t), \dot{\bar{x}}(t))$ (当 $\varepsilon \downarrow 0$ 时) 的 $(\widetilde{x}_\varepsilon, \widetilde{v}_\varepsilon) \in \operatorname{gph} F(t)$ 时成立. 于是在情形 (ii) 中得到近似伴随包含

$$\dot{p}_\varepsilon(t) \in \operatorname{co}\left\{u \in X^* \middle| (u, p_\varepsilon(t)) \in N((\widetilde{x}_\varepsilon, \widetilde{v}_\varepsilon); \operatorname{gph} F(t)) + \sqrt{\varepsilon}(0, \mathbb{B}^*)\right\}.$$

为了导出问题 $(P_{\rm M})$ 的广义 Euler-Lagrange 包含 (6.49), 需要对情形 (i) 和 (ii) 中关于 $p_\varepsilon(\cdot)$ 的近似伴随包含取极限 (当 $\varepsilon \downarrow 0$ 时), 因为这两个近似伴随包含相似, 为了明确起见, 只考虑第一种情形. 由逐点收敛 $(x_\varepsilon(t), \dot{x}_\varepsilon(t)) \to (\bar{x}(t), \dot{\bar{x}}(t))$ 和定理 3.60 中的基本法锥的鲁棒性质 (该性质成立是由于 F 的 SNC 假设), 注意到

$$\operatorname*{Limsup}_{\varepsilon \downarrow 0} N((x_\varepsilon(t), \dot{x}_\varepsilon(t)); \operatorname{gph} F(t)) = N((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F(t)).$$

也注意到, $p_\varepsilon(\cdot)$ 的近似伴随包含可等价地由 F 的基本上导数描述. 于是, 由定理中的强上导数正规性假设, $p_\varepsilon(\cdot)$ 的近似伴随包含则可由混合上导数 $D_M^* F$ 等价地描述. 利用定理 1.44 中 Lipschitz 连续性的混合上导数条件及上面的经典 Dunford 和 Mazur 定理, 类似定理 6.21 的证明可得到 (6.49).

下面考虑具有 (6.56) 中费用函数 φ_ε 的问题 (6.55) 中 $p_\varepsilon(b)$ 的横截性包含. 在该情形中, 利用定理 6.22 中的横截性条件 (6.45), 即可得 (6.45) 中的第一项, 其中 $\lambda = 1$ 并且 $\varphi(x_a, x_b) = \varphi_\varepsilon(x_b)$. 关键的条件

$$\operatorname{dist}\Big((x_\varepsilon(b), \beta_\varepsilon); \operatorname{epi} \varphi_\Omega\Big) > 0$$

确保了 $(x_\varepsilon(b), \beta_\varepsilon) \notin \operatorname{epi} \varphi_\Omega$ 对所有充分小的 $\varepsilon > 0$ 成立. 再利用定理 1.105 /推论 1.106, 对某 $\lambda_\varepsilon \geqslant 0$ 有

$$(-p_\varepsilon(b), -\lambda_\varepsilon) \in \bigcup_{(x,b) \in \prod((x_\varepsilon, \beta_\varepsilon); \operatorname{epi} \varphi_\Omega)} N((x, \beta); \operatorname{epi} \varphi_\Omega).$$

此外, 由于 φ_ε 在点 $\bar{x}(b)$ 的 SNEC 性质并因此在该点附近是 SNEC 的, 可以令 $\lambda_\varepsilon + \|p_\varepsilon(b)\| = 1$; 参见注 1.27(ii). 通过取极限 (当 $\varepsilon \downarrow 0$ 时) 并且考虑到定理 3.60 中的鲁棒稳定性结果, 可得到所需的横截性包含 (6.51) 并且 $\lambda \geqslant 0$. 由 φ_Ω 的 SNEC 性质 (该性质当 Ω 在点 $\bar{x}(b)$ 是 SNC 和 φ 在该点附近是 Lipschitz 连续时显然成立; 这是定理 3.90 的一个简单推论, 它甚至确保了 φ 在点 $\bar{x}(b)$ 的强 SNC 性质), 非平凡条件 $\lambda + \|p(b)\| = 1$ 可由关于 $(\lambda_\varepsilon, p_\varepsilon(b))$ 的非平凡条件通过取极限 (当 $\varepsilon \downarrow 0$ 时) 得到. 当 φ 在 $\bar{x}(b)$ 附近关于 Ω 是局部 Lipschitz 时, 由引理 5.23 可得横截性包含 (6.51) 和 (6.52) 是等价的. 注意到, 对 Lipschitz 连续费用函数, 包含 (6.5) 进一步蕴涵

$$-p(b) \in \lambda \partial \varphi(\bar{x}(b)) + N(\bar{x}(b); \Omega).$$

上面对无松弛问题 (P_M) 的任意中间局部极小点证明了定理中的广义 Euler-Lagrange 和横截性条件. 在这个一般的非凸情形中, 广义 Euler-Lagrange 包含 (6.49) 并不自动地蕴涵最大值条件 (6.50). 为了建立最大值条件 (6.50) 以及 (6.51), 遵循 Vinter [1289] 中定理 7.4.1 的证明, 它涉及有限维空间非凸微分包含的 Mayer 型问题 (P_M). 该定理的证明把非凸微分包含的约束 Mayer 问题化为无约束 Bolza (有限 Lagrange 函数) 问题, 进而化为一个可用直接方法导出最大值原理的光滑动态最优控制问题; 也请比照 6.3 节. 可验证上面发展的无穷维变分分析工具和所作的假设可将该证明扩展到自反和可分空间的情形. 这样一来就建立起了最大值条件 (6.50). 证毕. △

注 6.28(弱化假设下非凸微分包含的必要条件) 定理 6.27 中的一些假设, 尤其是那些在 (a)~(c) 中关于 Kadec 范数以及弱闭图像和上图的假设, 在对证明作适当修改后可放宽. 这涉及将定理 6.22 中的必要最优性条件应用到无约束 Bolza 问题 (6.55) 中. 该条件由基本/极限构造表述并且要求将推论 1.106 中的投影结果有效地用来估计距离函数在设集合外点的基本次梯度. 为了避免这些额外要求, 首先可将定理 6.27 的模糊离散逼近版本应用到无约束问题 (6.55) 中. 如定理 6.21 的证明, 这涉及 Fréchet法向量和次梯度. 然后当 $N \to \infty$ 和 $\varepsilon \downarrow 0$ 时取极限, 其中用相应于定理 1.103 中距离函数的 Fréchet次梯度来代替推论 1.106 中距离函数的基本次梯度. 不过这样做要复杂得多.

当 X 是有限维时, F 的 SNC 和强上导数正规性性质自动成立. 它也蕴涵着定理 6.27 中扩张端点函数 φ_Ω 的 SNEC 性质. 更进一步, 如果费用函数是局部 Lipschitz 的并且端点约束集通过由局部 Lipschitz 函数定义的有限多个等式和不等式给出, 那么在一般无穷维空间的情形, SNEC 性质是不需要的 (实际上, 在 Mangasarian-Fromovitz 型的规范条件下, 该性质自动成立).

推论 6.29 (具有等式和不等式约束微分包含的横截性条件) 设 $\bar{x}(\cdot)$ 为具有如下端点约束集的 Mayer 问题 (P_M) 的一个中间局部极小点:

$$\Omega := \Big\{x \in X \mid \varphi_i(x) \leqslant 0,\ i = 1, \cdots, m;\ \varphi_i(x) = 0,\ i = m+1, \cdots, m+r\Big\},$$

其中每一个 φ_i 以及费用函数 $\varphi_0 := \varphi$ 在 $\bar{x}(b)$ 附近是 Lipschitz 的. 假设除了扩张端点函数 φ_Ω 的 SNEC 性质外, 定理 6.27 的所有假设成立. 那么存在非负乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$ 和绝对连续伴随轨线 $p\colon [a, b] \to X^*$, 满足 Euler-Lagrange 和最大值条件 (6.49) 和 (6.50) 以及互补松弛条件

$$\lambda_i \varphi_i(\bar{x}(b)) = 0, \quad i = 1, \cdots, m$$

和横截性包含

$$-p(b) \in \sum_{i=0}^{m} \lambda_i \partial\varphi_i(\bar{x}(b)) + \sum_{i=m+1}^{m+r} \lambda_i \left[\partial\varphi_i(\bar{x}(b)) \cup \partial(-\varphi_i)(\bar{x}(b))\right].$$

更进一步, 如果所有 φ_i, $i = 0, \cdots, m+r$ 在 $\bar{x}(b)$ 处都是严格可微的, 那么存在乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$, 使得 $\lambda_i \geqslant 0$, $i = 0, \cdots, m$ 和伴随轨线 $p\colon [a,b] \to X^*$ 满足

$$-p(b) = \sum_{i=0}^{m+r} \lambda_i \nabla\varphi_i(\bar{x}(b))$$

以及上面的 Euler-Lagrange, Weierstrass-Pontryagin 和互补松弛条件.

证明　由 (6.52) 在 $\lambda := \lambda_0$ 的情形有

$$-p(b) \in \lambda_0 \partial\varphi_0(\bar{x}(b)) + N(\bar{x}(b); \Omega).$$

此外, 如果 Ω 在 $\bar{x}(b)$ 是 SNC 的, 那么 φ_Ω 在该点是 SNEC 的; 参见推论 3.89. 那么类似推论 6.24 的证明可完成推论的证明. 证毕. △

6.2.2　讨论和例子

在这一小节中考虑上面所得结果的某些推广和变形, 讨论一些相互之间的关联和例子. 首先注意到, 第 3 章中发展的完备的广义微分和 SNC 分析法则在由

$$x(b) \in F^{-1}(\Theta) \cap \Omega$$

(其中 $F\colon X \rightrightarrows Y$ 和 $\Theta \subset Y$) 给出的算子端点约束的情形中, 可用来导出定理 6.27 的各种推论和推广; 请比照 5.1 节中的数学规划问题. 下面对其他一些关联到微分包含必要最优性条件的重要问题作更详细的讨论.

注 6.30(上次微分横截性条件)　除了定理 6.21 中的假设之外, 设空间 X 具有 $\mathcal{C}^1$ Lipschitz 阻尼函数. 在定理 6.22 和定理 6.27 中 X 自反性的假设下, 这自动成立. 然后利用 6.1 和 6.2 节中的结果和定理 1.88(ii) 中 Fréchet次梯度的光滑变分描述, 可导出问题 (P) 和 (P_{M}) 及其相应的离散时间问题的必要最优性条件和通过描述目标和不等式约束函数的上次梯度来表述的横截性关系. 这可通过把它们化为描述目标和不等式约束的光滑函数情形来办到; 请比照定理 5.19 对于不可导规划的证明. 特别地, 在微分包含 (6.48) 的绝对连续轨道 $x\colon [a,b] \to X$ 上, 在定理 6.27 中关于 F 和 X 的假设下 (对 φ_i 不作任何假设), 考虑 Mayer 问题

$$\min \quad \varphi_0(x(b)) \quad \text{s.t.} \quad \varphi_i(x(b)) \leqslant 0, \quad i = 1, \cdots, m.$$

对于中间局部极小点 $\bar{x}(\cdot)$, 有如下的必要最优性条件: 对于 Fréchet上次梯度的每一个集合, $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x}(b))$, $i = 0, \cdots, m$, 存在乘子

$$(\lambda_0, \cdots, \lambda_m) \neq 0 \quad 使得 \quad \lambda_i \geqslant 0,\ i = 0, \cdots, m$$

以及绝对连续映射 $p\colon [a,b] \to X^*$ 满足 Euler-Lagrange 和最大值条件 (6.49) 和 (6.50) 并且

$$\lambda_i \varphi_i(\bar{x}(b)) = 0, \quad i = 1, \cdots, m \quad 和$$

$$p(b) + \sum_{i=0}^{m} \lambda_i x_i^* = 0.$$

要证明这些条件, 由上面的讨论, 剩下只需验证定理 6.27 中扩张端点函数 φ_Ω 的 SNEC 性质, 其中

$$\Omega := \{x \in X \mid \varphi_i(x) \leqslant 0, i = 1, \cdots, m\},$$

数据 φ, φ_i 是光滑的. 推论 3.87 可确保非线性规划中经典约束集的 SNC 性质; 请比照推论 6.24 和 6.29 的证明.

注 6.31(多目标控制问题的必要最优性条件) 上面发展的方法和结果可扩展到由微分包含控制的多目标最优化问题中. 给定映射 $f\colon X \to Z$ 和 Banach 空间中的子集 $\Theta \subset Z$ 并且 $0 \in \Theta$, 考虑上面 Mayer 问题 (P_{M}) 相应的多目标问题, 其中 (6.48) 受限于 $x(b) \in \Omega$ 的轨道 $\bar{x}(\cdot)$ 的广义序 (f, Θ)- 最优性理解为: 存在一个序列 $\{z_k\} \subset Z$ 满足 $z_k \to 0$ $(k \to \infty)$, 使得对 $\bar{x}(\cdot)$ 的 $W^{1,1}([a,b];X)$- 邻域中的任何可行轨道 $x(\cdot)$ 有

$$f(x(b)) - f(\bar{x}(b)) \notin \Theta - z_k, \quad k \in \mathbb{N};$$

请比照定义 5.53 和 5.3.1 小节中相关的讨论. 设

$$\mathcal{E}(f, \Omega, \Theta) = \{(x, z) \in X \times Z \mid f(x) - z \in \Theta, x \in \Omega\}$$

为关于序集 Θ 的限制映射 $f_\Omega = f + \Delta(\cdot;\Omega)$ 的“广义上图”. 从上面 $\bar{x}(\cdot)$ 的 (f, Θ)- 最优性定义中选取序列 $z_k \to 0$, 定义函数

$$\theta_k(x) := \operatorname{dist}((x, f(\bar{x}) - z_k); \mathcal{E}(f, \Omega, \Theta)), \quad k \in \mathbb{N}.$$

类似定理 6.27 的证明并用序列 $\theta_k(x)$ 替换其中的 $\theta_\beta(x)$, 这样就可得到所考虑的多目标控制问题的必要最优性条件. 与定理 6.27 相比, 所不同的只是横截性关系. 具体地, 除了定理 6.27 中关于 X 和 F 的假设外, 设 Z 是 WCG 和 Asplund 的并且广义上图集 $\mathcal{E}(f, \Omega, \Theta)$ 在 $(\bar{x}, \bar{z})$ 附近是局部闭的和在该点是 SNC 的, 其中 $\bar{z} := f(\bar{x})$. 那么存在伴随轨线 $p\colon [a,b] \to X^*$ 和伴随向量 $z^* \in N(0;\Theta)$, 不同时为零, 满足广义 Euler-Lagrange 包含 (6.49), Weierstrass-Pontryagin 最大值条件 (6.50) 和横截性包含

$$(-p(b), -z^*) \in N((\bar{x}(b), \bar{z}); \mathcal{E}(f, \Omega, \Theta)).$$

如果映射 f 在 $\bar{x}$ 附近相对于 Ω 是 Lipschitz 连续的和强上导数正规的, Ω 和 Θ 在点 $\bar{x}$ 和 0 附近分别是局部闭的, 那么由引理 5.23, 后者等价于

$$-p(b) \in \partial \langle z^*, f_\Omega \rangle (\bar{x}), \quad z^* \in N(0; \Theta).$$

注意到, 对上面类型的多目标控制问题在闭偏好关系的情形可类似地处理; 请比照 5.3.4 小节. 这样也可导出由微分包含控制的具有均衡约束的多目标 (以及 Mayer 和 Bolza 类型) 控制问题的必要最优性条件, 它们是 5.2 节和 5.3.5 小节中研究的 MPEC 和 EPEC 问题的相应动态问题.

注 6.32(Hamilton 包含)　当 $X = \mathbb{R}^n$ 时, 可得到问题 (P_{M})(以及问题 (P) 和前面注中讨论过的相应问题) 的松弛中间局部极小点的额外最优性条件, 该最优性条件是通过 Hamilton 函数

$$\mathcal{H}(x,p,t) := \sup\{\langle p, v\rangle \mid v \in F(x,t)\}$$

的基本次梯度表述的. 如果 F 是凸值的并且在 $(\bar{x}, \bar{v})$ 附近满足一些在 (H1) 中关于 F 的假设下自动成立的要求, 那么由 Rockafellar 对偶化定理 [1162, 定理 3.3] 有

$$\mathrm{co}\left\{u \in \mathbb{R}^n \middle| (u,p) \in N((\bar{x},\bar{v}); \mathrm{gph}\, F)\right\} = \mathrm{co}\left\{u \in \mathbb{R}^n \middle| (-u, \bar{v}) \in \partial \mathcal{H}(\bar{x}, p)\right\}.$$

这里对 t 的依赖并不重要, 因此忽略. 该对偶化关系的证明本质是有限维空间的; 也请比照 Ioffe[604, 定理 4] 和 Vinter[1289, 定理 7.6.5] 的证明. 因为凸化包含 (6.18) 的 Hamilton 函数和原来的 $\mathcal{H}(x,p,t)$ 显然是一致的, 由上面的对偶性关系可得, 定理 6.27 中的 Euler-Lagrange 包含 (6.49) 蕴涵了增广 Hamilton 包含

$$\dot{p}(t) \in \mathrm{co}\left\{u \in \mathbb{R}^n \middle| (-u, \dot{\bar{x}}(t)) \in \partial \mathcal{H}(\bar{x}(t), p(t), t)\right\} \quad \text{a.e. } t \in [a,b], \tag{6.58}$$

它是有限维状态空间中松弛极小点的一个必要最优性条件. 此外, 对于具有凸速率集 $F(x,t)$ 的问题 (P_{M}), Euler-Lagrange 包含 (6.49) 和 Hamilton 包含 (6.58) 等价. 注意到, (6.58) 是涉及基本次微分 $\partial\mathcal{H}(\bar{x}(t), p(t), t)$ 偏凸化的精细化 Hamilton 包含, 它明显地替换了涉及关于 (x,p) 的 Hamilton 函数的 Clarke 广义梯度 $\partial_C \mathcal{H}(\bar{x}(t), p(t), t) = \mathrm{co}\, \partial \mathcal{H}(\bar{x}(t), p(t), t)$ 的完全凸化版本

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \mathrm{co}\, \partial \mathcal{H}(\bar{x}(t), p(t), t) \quad \text{a.e. } t \in [a,b]. \tag{6.59}$$

值得注意的是, Hamilton 包含 (6.58) 和 (6.59) 关于 $F(x,t)$ 的凸化是不变的, 这与广义 Euler-Lagrange 包含 (6.49) 是不同的.

注 6.33(局部可控性)　前面小节中对必要最优性条件发展的方法也可用来研究有限维空间中关联到所谓的非凸微分包含的局部可控性问题. 给定 $x_0 \in X$, 记

$\mathcal{R}(x_0)$ 为微分包含 (6.48) 的可达集, 它是所有 $z \in X$ 组成的集合使得对 (6.48) 的某可行轨线 $x\colon [a,b] \to X$ 有 $x(b) = z$. 局部可控性, 在某种广泛的意义下, 是要导出微分包含 (6.48) 边界轨道的有效条件. 更为确切地, 考虑 $\bar{x}(b)$ 附近的局部 Lipschitz 映射 $g\colon X \to X$ 和 (6.48) 的轨道 $\bar{x}\colon [a,b] \to X$ 使得 $g(\bar{x}(b)) \in \mathrm{bd}\mathcal{R}(x_0)$. 那么除了 (H1) 和 (H2′) 外, 假设 $X = \mathbb{R}^n$, 则可找到向量 $x^* \in \mathbb{R}^*$ 使得 $\|x^*\| = 1$ 并且有伴随轨线 $p(\cdot)$ 满足具有边界/横截性条件

$$-p(b) \in \partial \langle x^*, g\rangle (\bar{x}(b)) \tag{6.60}$$

和 Weierstrass-Pontryagin 最大值条件 (6.50) 的广义 Euler-Lagrange 包含 (6.49). 此外, 如果可达集 $\mathcal{R}(x_0)$ 在 $\bar{x}(b)$ 附近是局部闭的, 那么也满足增广 Hamilton 包含 (6.48).

为了证明具有新横截性条件 (6.60) 的 Euler-Lagrange 最大值条件 (6.49) 和 (6.50), 遵循定理 6.27 的证明, 给定任何 $\varepsilon > 0$, 找到向量 $c_\varepsilon \in \mathbb{R}^n$ 和 (6.48) 的轨道 x_ε 使得 $\|g(x_\varepsilon(b)) - c_\varepsilon\| > 0$,

$$c_\varepsilon \to g(\bar{x}(b)), \quad x_\varepsilon(\cdot) \to \bar{x}(\cdot) \quad (\text{当 } \varepsilon \downarrow 0 \text{ 时在 } W^{1,1}([a,b];\mathbb{R}^n) \text{ 中的强收敛})$$

并且 $x_\varepsilon(\cdot)$ 是具有被积函数 (6.57) 和端点函数 $\varphi_\varepsilon := \|g(z) - c_\varepsilon\|$ 的问题 (6.55) 的无条件强局部极小点. 然后与定理 6.27 的证明一样, 唯一不同的是现在需要计算新函数 $\varphi_\varepsilon(\cdot)$ 在满足 $\|g(x_\varepsilon(b)) - c_\varepsilon\| > 0$ 的点 $\bar{x}_\varepsilon(b)$ 的基本次微分. 利用推论 3.43 中的次微分链式法则, 然后取极限 (当 $\varepsilon \downarrow 0$ 时), 与此同时考虑到 $\mathbb{R}^n$ 中单位球面的紧性, 可得到横截性条件 (6.60), 它是 (6.49) 和 (6.50) 的补充. 为了证明增广 Hamilton 包含 (6.58), 注意到, 所作的假设确保了凸化微分包含

$$\dot{x}(t) \in \mathrm{co}\, F(x(t),t) \quad \text{a.e. } t \in [a,b], \quad x(a) = x_0$$

可达集 $\widetilde{\mathcal{R}}(x_0)$ 的闭性和 $\mathcal{R}(x_0)$ 在 $\widetilde{\mathcal{R}}(x_0)$ 中的稠密性; 请比照定理 6.11. 因此由 $\mathcal{R}(x_0)$ 的局部闭假设可得 $\bar{x}(b)$ 是 $\widetilde{\mathcal{R}}(x_0)$ 的边界点, 于是 (6.58) 可由注 6.32 的讨论得到.

注意到, 状态空间 X 的有限维数在上面局部可控性的证明中是必须的, 它保证了在 X^* 的弱 * 拓扑下对偶单位球面的紧性; 由基本的 Josefson-Nissenzweig 定理, 这在无穷维空间中永远不成立. 与定理 6.27 的证明中无穷维的情形的差异是由这样的事实造成的: 在定理 6.27 的证明中, 实际上对集合 $\mathcal{R}(x_0) \times \{\varphi(\bar{x}(b))\}$ 和 $\mathrm{epi}\,\varphi_0$ (用定理 6.27 中的记号) 构成的局部集值系统利用了确切极值原理并且在极值系统中对第二个集合作 SNC 假设. 而在局部可控性的情形中, 则处理由集合 $\mathcal{R}(x_0)$ 和 $\{\bar{x}(b)\}$ 构成的局部集值系统, 其中第二个集合是单点集, 在无穷维空间中永不是

SNC 的. 另一方面, 在定理 6.27 的证明中和局部可控性的框架下并没有探索在可达集 $\mathcal{R}(x_0)$ 上作 SNC 要求的可能性, 这种要求可能引出其他假设来确保无穷维空间中满足必要最优性和局部可控性条件; 参见注 6.25(i) 中的结果和讨论.

作为该节的结束, 下面给出一些例子来阐明所得的结果和它们之间的关系. 首先说明偏凸化在广义 Euler-Lagrange 和 Hamilton 包含 (6.49) 和 (6.58) 中是不可避免的.

例 6.34(偏凸化在广义 Euler-Lagrange 和 Hamilton 最优性条件中的本质重要性) 在无端点约束的凸值微分包含的绝对连续轨道上, 存在一个二维的极小化线性函数的 Mayer 问题, 使得作为必要最优性条件的 Euler-Lagrange 包含 (6.49) 和 Hamilton 包含 (6.58) 的无 (偏) 凸化的类似结果不成立.

证明 对于凸值微分包含, 考虑下列 Mayer 问题, 其中 $x = (x_1, x_2) \in \mathbb{R}^2$:

$$\begin{cases} \min & J[x] := x_2(1) \\ \text{s.t.} & \dot{x}_1 \in [-\nu, \nu], \quad x_1(0) = 0, \\ & \dot{x}_2 = |x_1|, \quad x_2(0) = 0, \\ & \text{对} \quad \text{a.e. } t \in [0,1] \quad \text{和某} \quad \nu > 0 \quad \text{成立.} \end{cases}$$

易见 $\bar{x}(t) \equiv 0$ 是该问题的唯一最优解, 并且对于伴随轨线 $(p(t), -1) \in \mathbb{R}^2$ 的 Euler-Lagrange 包含 (6.49) 的无“co”的相应结果沿着该 $\bar{x}(\cdot)$ 给出关系式

$$\dot{p}(t) \in \{-1, 1\} \quad \text{a.e. } t \in [0,1]$$

和横截性条件 $p(1) = 0$. 更进一步, 由定理 1.34, 在此情形中 Euler-Lagrange 包含蕴涵具有如下形式的最大值条件

$$\langle p(t), \dot{\bar{x}}(t) \rangle = \max_{v \in [-\nu, \nu]} \langle p(t), v \rangle \quad \text{a.e. } t \in [0,1],$$

这给出 $p(t) \equiv 0$, 矛盾. 因为 $\mathcal{H}(p, x) = \nu \operatorname{sign} p - |x_1|$, (6.58) 中无“co”的 Hamilton 包含

$$(-\dot{p}(t), \dot{\bar{x}}(t)) \in \partial \mathcal{H}(\bar{x}(t), p(t)) \quad \text{a.e. } t \in [0,1]$$

导致如上的相同关系. 因此, 作为必要最优性条件是不成立的. △

下面两个例子阐明了广义 Euler-Lagrange 包含 (6.49) 和具有 (完全) 凸化 Hamilton 包含 (6.59) 的增广 Hamilton 包含 (6.58) 之间的关系.

例 6.35(广义 Euler-Lagrange 包含严格好于增广 Hamilton 包含) 存在紧值和凸值的多值函数 $F: \mathbb{R}^2 \rightrightarrows \mathbb{R}^2$, 它在 $\mathbb{R}^2$ 上是 Lipschitz 连续的并且使得对平面中的某些点 x, v, w, p 有

$$(-w, v) \in \operatorname{co} \partial \mathcal{H}(x, p) \quad \text{但} \quad w \notin \operatorname{co}\{u \in \mathbb{R}^2 \mid u \in D^* F(x, v)(-p)\}.$$

证明 定义 $F:\mathbb{R}^2 \rightrightarrows \mathbb{R}^2$ 为

$$F(x_1,x_2) := \{(\tau, \tau|x_1|+\nu) \in \mathbb{R}^2 \mid \tau \in [-1,1],\ \nu \in [0,\mu]\} \quad (\text{对某}\mu > 0),$$

其中对所有 $x=(x_1,x_2)\in\mathbb{R}^2$, 集合 $F(x)$ 是平面中的平行四边形. 相应的 Hamilton 函数为

$$\mathcal{H}(x_1,x_2,p_1,p_2) = |p_1 + p_2|x_1|| + \max\{p_2, 0\}.$$

考虑点 $x=(0,0)$, $v=(0,0)$ 和 $p=(0,-1)$. 注意到, 相应的集合 $F(x)$ 是矩形 $[-1,1]\times[0,\mu]$ 并且 p 是该集合在边界点 v 的外法向量. 该例子的主要特点是在 v 点支撑集合 $F(x)$ 的超平面 $x_2=0$ 与该集合的交多于一个点. 换言之, $\langle p, v\rangle$ 在 $v\in F(x)$ 上的最大值在无穷多个点达到. $\mathcal{H}$ 在点 $(0,0,0,-1)$ 的基本次微分及其凸化 (Clarke 广义梯度) 实际上在例 2.49 中已计算过. 因此

$$\operatorname{co}\partial\mathcal{H}(0,0,0,-1) = [-1,1]\times\{0\}\times[-1,1]\times\{0\} \subset \mathbb{R}^4.$$

取 $w=(-1,0)$, 有 $(-w,v)\in \operatorname{co}\partial\mathcal{H}(0,0,0,-1)$. 现在证明

$$(w,p) = (-1,0,0,-1) \notin \operatorname{cl}\operatorname{co} N((x,v);\operatorname{gph} F),$$

这就证明了该例子中的断言.

为此, 不计坐标的排列, F 的图像可表示为 $\operatorname{gph} F = E\times\mathbb{R}$, 其中

$$E := \big\{(x_1, \tau, |x_1|\tau + \nu) \in \mathbb{R}^3 \big| \tau\in[-1,1],\ \nu\in[0,\mu]\big\},$$

这里在 $(0,0,0)$ 点附近, 集合 E 和定义为 $\varphi(y,\tau) := \tau|y|$ 的 Lipschitz 函数 $\varphi:\mathbb{R}^2\to\mathbb{R}$ 的上图显然是一致的. 易见

$$\operatorname{co}\partial\varphi(0,0) = \partial\varphi(0,0) = \{(0,0)\}.$$

因此可计算

$$N((0,0,\varphi(0,0));\operatorname{epi}\varphi) = \bigcup_{\lambda\geqslant 0}\lambda[\partial\varphi(0,0)\times\{-1\}] = \big\{(0,0)\big\}\times(-\infty,0].$$

于是有

$$\operatorname{cl}\operatorname{co} N((0,0,0,0);\operatorname{gph} F) = \big\{(0,0,0)\big\}\times(-\infty,0].$$

特别地, 后面的锥不包含点 $(w,p)=(-1,0,0,-1)$, 但 $(-w,v)\in\operatorname{co}\partial\mathcal{H}(x,p)$. 证毕. △

下面的例子表明, 在凸值和非凸值微分包含的情形, 增广/精细化的 Hamilton 条件 (6.58) 可严格取代完全凸化的 (6.59).

例 6.36(偏凸化的 Hamilton 条件严格地改进其相应的完全凸化) 存在集值映射 $F:\mathbb{R}^n \rightrightarrows \mathbb{R}^n$, 它具有形式 $F(x)=g(x)S$, 其中 $S\subset\mathbb{R}^n$ 是紧的并且对每个 x, $g(x)$ 是 $\mathbb{R}^n$ 的一个连续依赖于 x 的线性同构, 使得对某 $(\bar{x},\bar{v},\bar{p})$ 有

$$\operatorname{co}\left\{u\in\mathbb{R}^n\Big|(u,\bar{v})\in\partial\mathcal{H}(\bar{x},\bar{p})\right\}\neq\left\{u\in\mathbb{R}^n\Big|(u,\bar{v})\in\operatorname{co}\partial\mathcal{H}(\bar{x},\bar{p})\right\}.$$

证明 如果 F 以上面的形式给定, 那么它的 Hamilton 函数可计算为

$$\mathcal{H}(x,p)=\sup\left\{\langle p,v\rangle\Big|v\in g(x)S\right\}=\sup\left\{\langle p,g(x)s\rangle\Big|s\in S\right\}=:\delta^*(g^*(x)p;S),$$

其中 $\delta^*(\cdot;S)$ 表示集合 S 的标准支撑函数. 因为 S 有界, 所以它的支撑函数是连续的. 记

$$\psi_s(x,p):=\langle s,g^*(x)p\rangle=\langle g(x)s,p\rangle$$

并假设 $g(\cdot)$ 是 Lipschitz 的. 利用标量化公式并考虑到 ψ 的结构, 在任何给定点 $(\bar{x},\bar{p})$ 有

$$\partial\mathcal{H}(\bar{x},\bar{p})=\bigcup_{s\in\partial\delta^*(0;S)}\partial\psi_s(\bar{x},\bar{p}).$$

由 ψ 关于 p 的线性性有

$$\partial\psi_s(\bar{x},\bar{p})=(\partial_x\psi_s(\bar{x},\bar{p}),g(\bar{x})s).$$

因此, 包含 $(u,0)\in\partial\psi_s(\bar{x},\bar{p})$ 蕴涵了 $s=0$, 因而 $u=0$.

根据上面的讨论, 需要找到集合 S, $\mathbb{R}^n$ 的 Lipschitz 连续线性同构族 $g(x)$ 以及点 $(\bar{x},\bar{p})\in\mathbb{R}^n\times\mathbb{R}^n$, 使得 $0\in S$ 和 $\operatorname{co}\partial\mathcal{H}(\bar{x},\bar{p})$ 包含一个点对 $(u,0)$, 其中 $u\neq 0$. 特别地, 对于 $n=2$, 可如下得到. 设

$$S:=\{(y_1,y_2)\in\mathbb{R}^2\mid |y_1|\leqslant 1,\ y_2=0\},\quad g^*(x):=\begin{pmatrix}1 & |x_1|\\ 1 & 1\end{pmatrix},$$

$\bar{x}:=(0,0)$ 和 $\bar{p}:=(0,1)$. 那么

$$\delta^*((w_1,w_2);S)=w_1\quad 和\quad \mathcal{H}(x,p)=|p_1+p_2|x_1||.$$

直接计算 (请比照例 2.49) 可得集合 $\operatorname{co}\mathcal{H}(\bar{x},\bar{p})$ 是下面四个点的凸包: $(1,0,1,0)$, $(-1,0,-1,0)$, $(1,0,-1,0)$ 和 $(-1,0,1,0)$. 因此

$$\left\{u\in\mathbb{R}\Big|(u,0)\in\operatorname{co}\partial\mathcal{H}(\bar{x},\bar{p})\right\}=[-1,1].$$

证毕. △

6.3 具有光滑动态的连续时间系统的最大值原理

在这一节研究由无穷维空间中微分方程控制的最优控制问题如下, 它显式地涉及约束控制输入 $u(\cdot)$:

$$\dot{x} = f(x, u, t), \quad u(t) \in U \quad \text{a.e. } t \in [a, b], \tag{6.61}$$

其中 $f : X \times U \times [a, b] \to X$, X 为 Banach 空间并且 U 为度量控制空间. 虽然这种类型的控制系统可化为微分包含 $\dot{x} \in F(x, t)$, 其中 $F(x, t) := f(x, U, t)$, 但由于 (6.61) 中的控制输入是显式给出的, 控制区域 U 不依赖于 x (它可能依赖于 t), 故可以考虑到其具体的特点而发展研究这类动态系统的有效方法.

在整节中, 假设系统 (6.61) 是光滑动态的, 这意味着速率映射 f 在所考虑的最优解附近关于状态变量 x 是连续可微 ($\mathcal{C}^1$) 的. 尽管作这样的假设, 但由于控制几何约束 $u(t) \in U$ (a.e. $t \in [a, b]$) 是由一般的控制集 U 定义的, 控制系统 (6.61) 和在其可行控制和轨道上的最优化问题内在地涉及非光滑性. 例如对于 $U = \{0, 1\}$ 的最简单/经典的最优控制问题便是如此.

本节主要的精力放在光滑动态系统 (6.61) 的 Mayer 类型控制问题上, 它受限于由具体在极小点仅仅是 Fréchet(可能不是严格) 可微函数的等式和不等式给出的有限多个端点约束. 在一般的 Banach 空间中, 除了关于 X 的自反性和可分性的假设以及定理 6.27 对 $F(x, t) = f(x, U, t)$ 所作的序列法紧性 (SNC) 和强上导数正规性的假设之外, 无需作别的假设. 对于这类问题, 目的是要导出具有 Pontryagin 最大值原理 (PMP) 形式的必要最优性条件. 与 6.1 和 6.2 节不同, 这里使用的技巧可追溯到最优控制理论中涉及最优控制的针形变分的经典方法. 对于不可导费用和不等式约束函数的情形, 也导出了最大值原理类型的加强结果, 其中以上次微分描述横截性条件. 得到这些条件不需对所论的状态空间加任何光滑性假设, 而这样的光滑性假设对上面的数学规划和动态优化问题的必要最优条件是需要的; 请比照定理 5.19 和注 6.30. 另外本节的结果本质地利用了光滑控制系统 (6.61) 及其端点作约束的特殊性质, 因此一般来说独立于小节 6.1 和 6.2 中得到的那些结果.

本节结构如下. 6.3.1 小节包括对主要假设和结果的阐述以及从 Fréchet可微端点函数情形的最大值原理导出具有上次微分横截性条件的最大值原理. 也讨论最大值原理在具有中间状态约束控制问题以及一些时滞系统中的可能扩展. 6.3.2 小节对 Banach 空间中自由端点控制问题证明了 PMP, 这比端点控制问题 PMP 的证明要简单得多. 6.3.3 小节处理涉及不等式类型端点约束最优控制问题. 最后, 在 6.3.4 小节中, 利用 Brouwer 不动点定理, 在仅在最优点可导的连续函数情形, 导出了横截性条件.

6.3.1　主要结果的阐述和讨论

本节阐述关于具有固定左端点 $x(a) = x_0$ 控制系统 (6.61) 的主要结果, 该情形实际上并没有太大地限制一般性, 但结果的描述和证明都更加简单和方便; 本节末讨论主要结果的各种扩展.

记 $\mathcal{A}$ 为有可测控制 $u(\cdot)$ 生成的容许控制轨道对 $\{u(\cdot), x(\cdot)\}$ 组成的集合, 满足逐点约束 $u(t) \in U$ (a.e. $t \in [a, b]$), 相应于 (6.61), $x(a) = x_0$ 的解 $x(\cdot)$ 定义为

$$x(t) = x_0 + \int_a^t f(x(s), u(s), s)\mathrm{d}s, \quad \forall\ t \in [a, b], \tag{6.62}$$

其中的积分理解为 Bochner 意义下的积分; 请比照定义 6.1. 众所周知, (6.62) 的任何解在 $[a, b]$ 上是绝对连续的. 此外, 如果 X 具有 Radon-Nikodým性质 (参见 6.1.1 小节, 这里的结果没有这个假设), 那么它在 $[a, b]$ 上是几乎处处可微的并且对几乎处处的 $t \in [a, b]$ 满足微分方程 (6.61). 这节需要的是积分表示 (6.62), 它被作为 Banach 空间中微分方程 (6.61) 可行解/轨线的定义.

在状态空间 X 中, 给定实值函数 φ_i, $i = 0, \cdots, m + r$, 所要研究的最优控制问题如下:

$$\min \quad J[u, x] = \varphi_0(x(b)), \quad (u, x) \in \mathcal{A}, \tag{6.63}$$

其中端点约束为

$$\varphi_i(x(b)) \leqslant 0, \quad i = 1, \cdots, m, \tag{6.64}$$

$$\varphi_i(x(b)) = 0, \quad i = m + 1, \cdots, m + r. \tag{6.65}$$

满足端点约束 (6.64) 和 (6.65) 的容许解 $(u, x) \in \mathcal{A}$ 称为问题 (6.63)~(6.65) 的可行解. 因而对具有自由端点问题, 即没有端点约束 (6.64) 和 (6.65) 的容许和可行解不作区别. 这里总假设 (6.63)~(6.65) 的可行解集是非空的.

称问题 (6.63)~(6.65) 的可行解 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 是最优的, 如果

$$J[\bar{u}, \bar{x}] \leqslant J[u, x] \quad (\forall\ (u, x) \in \mathcal{A})$$

并满足端点约束 (6.64) 和 (6.65). 对于问题的给定最优解 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$, 下面的目标是要导出 PMP 型的必要条件. 虽然所给的必要条件是针对 (整体) 最优解的, 但由下面的证明可看到, 只要 (u, x) 对于 (6.63)~(6.65) 是可行的并且 $\|x(t) - \bar{x}(t)\| < \varepsilon$ 对所有 $t \in [a, b]$ 和某 $\varepsilon > 0$ 成立, 那么所得的结果对局部极小点 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 在 $J[\bar{u}, \bar{x}] \leqslant J[u, x]$ 的意义下成立. 这对应于 6.1.2 小节中 $F(x, t) = f(x, U, t)$ 的强局部极小点.

给定 (6.63)~(6.65) 的最优解 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$, 在整节中作如下常驻假设:

— 状态空间 X 是 Banach 的;

— 控制集合 U 是可分 Banach 空间中的 Souslin 子集 (即 Borel 子集的连续像);

— 存在包含 $\bar{x}(t)$ 的开子集 $O \subset X$, 使得 f 关于 x 是 Fréchet可微的并且 $f(x,u,t)$ 和 $\nabla_x f(x,u,t)$ 关于 (x,u) 连续, 关于 t 可测, 对所有 $x \in O$, $u \in U$ 和 a.e. $t \in [a,b]$ 范数有界且由可和函数界定;

— 函数 φ_i $(i = m+1, \cdots, m+r)$ 在 $\bar{x}(b)$ 附近是连续的, 并且在该点是 Fréchet可微的.

注意到, 控制集合 U 可能以一般的可测方式依赖于 t, 这样一来就可以利用标准可测选择的结果; 参见文献 [54, 229, 1165] 以及其中的参考文献.

关于描述目标和不等式约束函数 φ_i $(i = 0, \cdots, m)$ 的适当假设将在定理 6.37 中给出. 注意到, 对它们的基本假设仅要求在 $\bar{x}(b)$ 的 Fréchet可微性 (甚至不是该点附近的连续性), 而上次微分条件对任意 Banach 空间中更为广泛的一类不可导函数成立.

为了阐述最大值原理之间的关系, 定义系统 (6.61) 的 Hamilton-Pontryagin 函数为

$$H(x,p,u,t) := \langle p, f(x,u,t) \rangle, \quad p \in X^*.$$

注意到, 6.2 节中定义的关于 $F(x,t) = f(x,U,t)$ 的 Hamilton 函数对应于在整个控制区域上函数 $H(x,p,u,t)$ 关于 u 的极大化

$$\mathcal{H}(x,p,t) = \max\{H(x,p,u,t) \mid u \in U\}.$$

也注意到, H 关于状态和伴随变量 (x,p) 是光滑的, 当然对 $\mathcal{H}$ 并非如此.

定理 6.37(光滑控制系统的最大值原理) 设 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 为问题 (6.63)~(6.65) 在所作常驻假设下的最优解. 假设函数 φ_i $(i = 0, \cdots, m)$ 在最优端点 $\bar{x}(b)$ Fréchet可微. 那么存在乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$ 满足

$$\lambda_i \geqslant 0, \quad i = 0, \cdots, m,$$

$$\lambda_i \varphi_i(\bar{x}(b)) = 0, \quad i = 1, \cdots, m,$$

并且使得下列最大值条件成立:

$$H(\bar{x}(t), p(t), \bar{u}(t), t) = \max_{u \in U} H(\bar{x}(t), p(t), u, t), \quad \text{a.e. } t \in [a,b], \tag{6.66}$$

其中绝对连续映射 $p : [a,b] \to X^*$ 是伴随系统

$$\dot{p} = -\nabla_x H(\bar{x}, p, \bar{u}, t), \quad \text{a.e. } t \in [a,b] \tag{6.67}$$

的轨道并且有横截性条件

$$p(b) = -\sum_{i=0}^{m+r} \lambda_i \nabla \varphi_i(\bar{x}(b)). \tag{6.68}$$

类似 (6.61), 系统 (6.67) 的解 (伴随轨道) 是在积分/适度意义下的, 即

$$p(t) = p(b) + \int_t^b \nabla_x H(\bar{x}(s), p(s), \bar{u}(t), s)\mathrm{d}s, \quad t \in [a, b],$$

其中 $\nabla_x H(\bar{x}, p, \bar{u}, t) = \langle p, \nabla_x f(\bar{x}, \bar{u}, t)\rangle$. 也注意到, 横截性条件 (6.68) 和推论 6.29 中一致. 然而, 现在端点函数没有假设为在点 $\bar{x}(b)$ 是严格可微的.

定理 6.37 的证明将在 6.3.2 和 6.3.4 小节中给出. 下面先阐述并证明该定理相应的上次微分情形, 这对不可微函数 φ_i $(i = 0, \cdots, m)$ 给出了横截性条件 (6.68) 的一个扩展, 它由定理 6.37 和 Fréchet次梯度的光滑变分描述得到.

定理 6.38(具有以 Fréchet上次梯度描述的横截性条件的最大值原理) 设 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 为最优控制问题 (6.63)–(6.65) 在所作的常驻假设下的最优解. 那么对每一个 Fréchet上次梯度 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x}(b))$, $i = 0, \cdots, m$, 存在乘子 $(\lambda_0, \cdots, \lambda_{m+r}) \neq 0$ 满足定理 6.37 中的符号和互补松弛条件, 使得最大值条件 (6.66) 成立, 其中伴随系统 (6.67) 相应的轨道 $p(\cdot)$ 满足横截性条件

$$p(b) + \sum_{i=0}^{m+r} \lambda_i x_i^* = 0. \tag{6.69}$$

证明 任取上次梯度 $x_i^* \in \widehat{\partial}^+\varphi_i(\bar{x}(b))$ $(i = 0, \cdots, m)$ 并利用定理 1.88 中关于 $-x_i^*$ 的光滑变分描述的断言 (i), 它对任何 Banach 空间都成立. 这样一来就可找到函数 $s_i : X \to \mathbb{R}$ 对 $i = 0, \cdots, m$ 满足关系

$$s_i(\bar{x}(b)) = \varphi_i(\bar{x}(b)), \quad \text{在 } \bar{x}(b) \text{ 附近 } s_i(x) \geqslant \varphi_i(x),$$

并且使得 $s_i(\cdot)$ 在点 $\bar{x}(b)$ Fréchet可微以及 $\nabla s_i(\bar{x}(b)) = x^*$, $i = 0, \cdots, m$. 由这些函数的构造容易推导出 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 是如下控制问题的一个最优解:

$$\min \quad \widetilde{J}[u, x] = s_0(x(b)) \quad (\text{在 } (u, x) \in \mathcal{A} \text{ 上})$$

且受限于不等式和等式端点约束

$$s_i(x(b)) \leqslant 0, \quad i = 1, \cdots, m$$

和 (6.65), 其中 $\mathcal{A}$ 是本小节开始时定义的可行控制轨道对所组成的集合. 该最优控制问题的初始数据满足定理 6.37 的所有假设. 因此, 对该问题利用上面的最大值原理并考虑到 $\nabla s_i(\bar{x}(b)) = x_i^*$, $i = 0, \cdots, m$, 这就完成了定理的证明. 证毕. △

可以观察到定理 6.38 的阐述和证明中涉及上次微分横截性条件的部分与定理 5.19 中数学规划里关于上次微分最优性条件的差异. 根据 Fréchet次梯度的光滑变分描述, 这两个结果都归结为它们相应的光滑情形. 在定理 5.19 的情形中, 需要费用和约束函数的连续可微性 (更为确切地, 严格可微性) 才能够应用光滑非线性规划中相应的必要条件. 由定理 1.88(ii), 这需要 Banach 空间几何上的额外假设才能确保 Fréchet次梯度的 $\mathcal{C}^1$ 刻画. 另一方面, 由定理 1.88(i) 的适当光滑变分刻画, 定理 6.38 依赖定理 6.37. 定理 6.37 仅要求端点函数在最优点的 Fréchet可微性. 注意到, 涉及最优控制问题的定理 6.37 和定理 6.38 对于 (6.61) 中取 $f=0$ 的情形显然蕴涵 5.1.3 小节中对于具有等式和不等式约束的数学规划问题结果的相应改进.

注 6.39(轨道在两个端点和中间点约束的控制问题) 在 6.3.2~6.3.4 小节中, 从定理 6.37 的证明中可以看到, 在对初始数据的同样假设下, 只需对该证明稍作修改就可对具有在 $t=a$ 和 $t=b$ 的 (6.64) 和 (6.65) 形式的端点约束以及具有依赖于 $x(a)$ 和 $x(b)$ 的费用函数 φ_0 的最优控制问题导出类似的必要最优性条件. 在此情形中, 绝对连续伴随轨线 $p:[a,b]\to X^*$ 上的横截性条件 (6.68) 被替换为

$$(p(a),-p(b))=\sum_{i=0}^{m+r}\lambda_i\nabla\varphi_i(\bar{x}(a),\bar{x}(b)).$$

更进一步, 可类似处理涉及中间状态约束, 即轨道的约束在区间中的点 $\tau_i\in[a,b]$ 给出. 例如, 考虑修改后的问题 (6.63)–(6.65) 使得

$$\varphi_i=\varphi_i(x(a),x(\tau),x(b)),\quad i=0,\cdots,m+r,$$

其中 $\tau\in[a,b]$ 是区间中的点. 那么定理 6.37 和修改后的状态约束问题的必要最优性条件的差异在于现在有不连续的伴随轨线 $p(\cdot)$ 且在中间点 $t=\tau$ 具有跳跃条件, 所有这些都体现在如下的横截性条件中:

$$(p(a),p(\tau+0)-p(\tau-0),-p(b))=\sum_{i=0}^{m+r}\lambda_i\nabla\varphi_i(\bar{x}(a),\bar{x}(\tau),\bar{x}(b)).$$

类似地, 在具有中间状态约束的情形, 可以修改定理 6.38 中的上次微分条件.

注 6.40(时滞控制系统的最大值原理) 定理 6.37 和定理 6.38 的结果可以推广到状态和控制变量具有时滞的各种系统. 例如, 考虑状态变量具有常数时滞 $\theta>0$ 的标准系统:

$$\begin{cases}\dot{x}(t)=f(x(t),x(t-\theta),u(t),t), & \text{a.e. } t\in[a,b],\\ x(t)=c(t), \quad t\in[a-\theta,a], & \\ u(t)\in U, \quad \text{a.e. } t\in[a,b]. & \end{cases}$$

该系统是在具有状态空间为 Banach 空间的可测控制和绝对连续轨道以及初始“尾”映射 $c\colon [a-\theta, a] \to X$ 之上,“尾”映射 c 对时滞过程的启动是必要的. 记 $\mathcal{A}$ 为满足上面时滞系统的容许对 $\{u(\cdot), x(\cdot)\}$ 的集合并且定义相应的 Hamilton-Pontryagin 函数为

$$H(x, y, p, u, t) := \langle p, f(x, y, u, t)\rangle, \quad p \in X^*,$$

其中 y 表示时滞变量 $x(t-\theta)$. 现在考虑问题 (6.63)–(6.65) 并且 $\mathcal{A}$ 表示时滞系统容许对的集合, 可得到相应于定理 6.37 和定理 6.38 的结果, 其中伴随系统为

$$-\dot{p}(t) = \begin{cases} \nabla_x H(x(t), x(t-\theta), p(t), u(t), t) \\ \qquad + \nabla_y H(x(t+\theta), x(t), p(t+\theta), u(t+\theta), t), \\ \qquad \text{a.e. } t \in [a, b-\theta]; \\ \nabla_x H(x(t), x(t-\theta), p(t), u(t), t), \quad \text{a.e. } t \in [b-\theta, b]. \end{cases}$$

实际上, 这些结果可以通过将 X 中的时滞控制系统约化为状态空间 X^N 中当 N 充分大时的无时滞控制系统来证明. 更进一步, 对于具有依赖于时间和状态变量的更一般的时滞以及具有时间分布时滞的控制问题, 定理 6.37 和定理 6.38 的证明中的方法可以用来导出类似的结果.

注 6.41(中立型泛函微分控制系统) 描述这类控制系统动态的微分方程中时滞不仅出现在状态变量中, 也出现在速率变量中. 在 $[a-\theta, a]$ 上具有适当初始条件的典型模型为

$$\dot{x}(t) = f(x(t), x(t-\theta), \dot{x}(t-\theta), u(t), t), \quad u(t) \in U, \quad \text{a.e. } t \in [a, b].$$

这类系统与前面注解中考虑的标准常微分方程控制系统和时滞系统有着本质的不同. 对于变分分析, 实质上它们更难并且呈现出许多上面考虑过的控制系统所不固有的现象; 读者可以在第 7 章的评论中找到更多的讨论, 在那里更详细地考虑这类系统及其扩展. 现在注意到, 虽然对具有 Banach 状态空间和凸速率集合 $f(x, y, z, U, t)$ 的情形, 由类似方法可以导出定理 6.37 和定理 6.38 形式的必要最优性条件, 但对中立型控制系统, 甚至在有限维空间中无端点约束的情形, 类似的 Pontryagin 最大值原理一般并不成立. 比如对如下二维控制问题:

$$\begin{cases} \min & J[u, x] = x_2(2) \\ \text{s.t.} & \dot{x}_1(t) = u(t), \quad \dot{x}_2(t) = \dot{x}_1^2(t-1) - u^2(t), \quad t \in [0, 2], \\ & x_1(t) = x_2(t) = 0, \quad t \in [-1, 0]; \quad |u(t)| \leqslant 1, \quad t \in [0, 2] \end{cases}$$

的最优控制

$$\bar{u}(t) = 0, \quad t \in [0, 1) \quad \text{和} \quad \bar{u}(t) = 1, \quad t \in [1, 2],$$

Pontryagin 最大值原理就不成立. 读者可在 Gabasov 与 Kirillova 的书 [485; 3.6 节] 中找到这个例子的完整计算; 类似的计算也可参见 6.4.6 小节中的例 6.70, 它是关于该问题的有限差分版本.

6.3.2 自由端点问题的最大值原理

本小节研究问题 (6.63), 其中 $\mathcal{A}$ 是具有左端点 $x(a)=x_0$ 的控制系统 (6.61) 的容许对 $\{u(\cdot),x(\cdot)\}$ 的集合; 更确切的阐述请参见 6.3.1 小节的开始. 尽管左端点总是固定的, 该问题称为最优控制的自由端点问题; 记住在容许轨道的右端没有约束 (6.64) 和 (6.65). 由下面的证明可知, 自由端点问题 (6.63) 和约束问题 (6.63)~(6.65) 有着本质的区别; 此外, 具有不等式和等式端点约束的问题之间也有着本质的区别. 无约束和约束问题的主要区别在于对 (6.63) 的情形, 所有的容许轨道都是可行的并且在变换容许控制 $u(\cdot)\in U$ 时, 不需要关心是否满足端点约束. 注意到, 在所考虑的问题中, 上面 (任意) 几何类型的控制约束总是存在的, 它们将最优控制问题和经典的变分法区分开来, 并且表明了最优控制中固有的内在非光滑性.

本小节对问题 (6.63) 在对给定数据 (U,X,f,φ_0) 所作的假设下, 给出定理 6.37 中最大值原理的证明. 注意到, 在此情形, 横截性条件 (6.68) 约化为

$$p(b)=-\nabla\varphi_0(\bar{x}(b)), \tag{6.70}$$

即在 (6.68) 中 $\lambda_0=1$ 和 $\lambda_i=0,\ i=1,\cdots,m+r$. 事实上, 在 (6.68) 中, 如果 $\lambda_0=0$ 和 $p(b)=0$, 那么由伴随系统 (6.67) 关于 p 的线性性, 对所有 $t\in[a,b]$ 有 $p(t)\equiv 0$, 这与定理 6.37 中的非平凡条件 $(p(\cdot),\lambda_0)\neq 0$ 矛盾.

对于自由端点问题 (6.63), 定理 6.37 的证明是纯分析的, 它不需要使用任何凸分离或类似的几何结果. 这与涉及 5.3.3 和 5.3.4 小节中给出的不等式和等式端点约束的定理 6.37 的证明是不同的. 问题 (6.63) 在定理 6.37 的证明中, 基本的工具是 (6.63) 中费用泛函的增量公式和利用最优控制中所谓的"针形"变分 (有时称为"McShane 变分").

下面从增量公式开始. 在 (6.62) 中给定两个容许控制 $\bar{u}(t),u(t)\in U$ (注意在回到定理 6.37 的证明之前, $\bar{u}(\cdot)$ 不必是最优的) 和相应的解 $\bar{x}(\cdot),x(\cdot)$. 记

$$\Delta\bar{u}(t):=u(t)-\bar{u}(t),\quad \Delta\bar{x}(t):=x(t)-\bar{x}(t),\quad \Delta J[\bar{u}]:=\varphi_0(x(b))-\varphi_0(\bar{x}(b)).$$

下面的意图是得到费用泛函增量 $\Delta J[\bar{u}]$ 的以 Hamilton-Pontryagin 函数给出的一个方便的表示, 其中的 Hamilton-Pontryagin 函数的值是关于容许对 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 和具有边界条件 (6.70) 的伴随系统 (6.67) 的对应轨道 $p(\cdot)$. 其中对所有该类的表达式都会用到一样的标准符号 $o(\cdot)$.

引理 6.42(费用泛函的增量公式)　利用如上定义的符号 Δ, 设

$$\Delta_u H(\bar{x}(t), p(t), \bar{u}(t), t) := H(\bar{x}(t), p(t), u(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t).$$

那么有

$$\Delta J[\bar{u}] = -\int_a^b \Delta_u H(\bar{x}(t), p(t), \bar{u}(t), t)\mathrm{d}t + \eta,$$

其中余项 η 为 $\eta = \eta_1 + \eta_2 + \eta_3$,

$$\eta_1 := o(\|\Delta\bar{x}(b)\|), \quad \eta_2 := -\int_a^b o(\|\Delta\bar{x}(t)\|)\mathrm{d}t,$$

$$\eta_3 := -\int_a^b \left\langle \frac{\partial \Delta_u H(\bar{x}(t), p(t), \bar{u}(t), t)}{\partial x}, \Delta\bar{x}(t) \right\rangle \mathrm{d}t.$$

证明　因为 φ_0 在点 $\bar{x}(b)$ 是 Fréchet可微的, 所以有表达式

$$\Delta J[\bar{u}] = \varphi_0(x(b)) - \varphi_0(\bar{x}(b)) = \langle \nabla\varphi_0(\bar{x}(b)), \Delta\bar{x}(b)\rangle + o(\|\Delta\bar{x}(b)\|).$$

考虑到状态和伴随方程的解满足 (由定义) Newton-Leibniz 公式, 并且对 Bochner 积分利用分步积分可得恒等式

$$\langle p(b), \Delta\bar{x}(b)\rangle = \int_a^b \langle \dot{p}(t), \Delta\bar{x}(t)\rangle \,\mathrm{d}t + \int_a^b \langle p(t), \Delta\dot{\bar{x}}(t)\rangle \,\mathrm{d}t,$$

其中 $p : [a, b] \to X^*$ 是解集中任意的绝对连续映射. 由 (6.70) 中关于 $p(b)$ 的边界条件有

$$\Delta J[\bar{u}] = -\int_a^b \langle \dot{p}(t), \Delta\bar{x}(t)\rangle \,\mathrm{d}t - \int_a^b \langle p(t), \Delta\dot{\bar{x}}(t)\rangle \,\mathrm{d}t + o(\|\Delta\bar{x}(b)\|).$$

接着变换上面的第二个积分项. 利用方程

$$\Delta\dot{\bar{x}}(t) = f(\bar{x}(t) + \Delta\bar{x}(t), \bar{u}(t) + \Delta\bar{u}(t), t) - f(\bar{x}(t), \bar{u}(t), t),$$

Hamilton-Pontryagin 函数 $H(x, p, u, t)$ 的定义和 f 关于 x 的光滑性有

$$\begin{aligned}
&\int_a^b \langle p(t), \Delta\dot{\bar{x}}(t)\rangle \,\mathrm{d}t \\
&= \int_a^b \big[H(\bar{x}(t) + \Delta\bar{x}(t), p(t), \bar{u}(t) + \Delta\bar{u}(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t)\big]\mathrm{d}t \\
&= \int_a^b \big[H(\bar{x}(t), p(t), \bar{u}(t) + \Delta\bar{u}(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t)\big]\mathrm{d}t \\
&\quad + \int_a^b \left\langle \frac{\partial \Delta_u H(\bar{x}(t), p(t), \bar{u}(t) + \Delta\bar{u}(t), t)}{\partial x}, \Delta\bar{x}(t) \right\rangle \mathrm{d}t + \int_a^b o(\|\Delta\bar{x}(t)\|)\mathrm{d}t.
\end{aligned}$$

最后, 考虑到 $p(\cdot)$ 是由 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 生成的伴随系统 (6.67) 的解, 即完成证明. △

在上面的增量公式中, 控制 $\bar{u}(\cdot)$ 和 $u(\cdot)$ 都是满足逐点控制约束的任意可测映射. 现在构造 $u(\cdot)$ 为参考控制 $\bar{u}(\cdot)$ 的一个特殊扰动, 称为该控制的针形变分, 或者单独针形变分. 也就是说, 固定任意数 $\tau\in[a,b)$ 和 $\varepsilon>0$ 使得 $\tau+\varepsilon<b$, 任取一点 $v\in U$, 构造容许控制 $u(t)$ $(t\in[a,b])$ 为

$$u(t):=\begin{cases} v, & t\in[\tau,\tau+\varepsilon),\\ \bar{u}(t), & t\notin[\tau,\tau+\varepsilon).\end{cases} \tag{6.71}$$

所得到的扰动控制与参考控制的差异仅在小的时间区间 $[\tau,\tau+\varepsilon)$ 上, 在该小区间上它取控制集 U 中的任意值; "针形变分"的称谓便来源于此. 对于依赖参数 (τ,ε,v) 的相应轨道增量 $\Delta\bar{x}(t)$, 显然有

$$\Delta\bar{x}(t)=0,\quad t\in[a,\tau].$$

接下来, 对 $t\in(\tau,b]$ 估计 $\Delta\bar{x}(t)$, 这在引理 6.4.3 中给出. 下面记 ℓ 为 $f(\cdot,v,t)$ 的一致 Lipschitz 常数, 它的存在性由所作的假设确定. 简单起见, 假设 ℓ 不依赖 t, 虽然所作的假设允许它依赖于 t 且在 $[a,b]$ 上是可和的, 但这并不改变所得的结果.

引理 6.43(针形变分下轨道的增量) 设 $\Delta\bar{x}(\cdot)$ 为 $\bar{x}(\cdot)$ 的相应于 $\bar{u}(\cdot)$ 具有参数 (τ,ε,v) 的针形变分 (6.71) 的增量. 那么存在不依赖 (τ,ε) (可能依赖 v) 的常数 $K>0$, 使得

$$\|\Delta\bar{x}(t)\|\leqslant K\varepsilon,\quad \forall\, t\in[a,b].$$

证明 因为 $\Delta\bar{x}(\tau)=0$, 由 (6.62) 有

$$\Delta\bar{x}(t)=\int_\tau^t\big[f(\bar{x}(s)+\Delta\bar{x}(s),v,s)-f(\bar{x}(s),\bar{u}(s),s)\big]\mathrm{d}s,\quad \tau\leqslant t\leqslant\tau+\varepsilon.$$

考虑到 f 关于 x 的具有 Lipschitz 常数 ℓ 的一致 Lipschitz 连续性, 并且记 $\Delta_v f(\bar{x}(s),\bar{u}(s),s):=f(\bar{x}(s),v,s)-f(\bar{x}(s),\bar{u}(s),s)$, 于是有

$$\begin{aligned}\|\Delta\bar{x}(t)\|&\leqslant\int_\tau^t\|f(\bar{x}(s)+\Delta\bar{x}(s),v,s)-f(\bar{x}(s),\bar{u}(s),s)\|\mathrm{d}s\\&\leqslant\int_\tau^t\|\Delta_v f(\bar{x}(s),\bar{u}(s),s)\|\mathrm{d}s+\ell\int_\tau^t\|\Delta\bar{x}(s)\|\mathrm{d}s.\end{aligned}$$

利用记号

$$\alpha(t):=\int_\tau^t\|\Delta_v f(\bar{x}(s),\bar{u}(s),s)\|\mathrm{d}s\quad 和\quad \beta(t):=\|\Delta\bar{x}(t)\|,$$

上面的估计可改写为

$$\beta(t) \leqslant \alpha(t) + \ell \int_\tau^t \beta(s)\mathrm{d}s, \quad \tau \leqslant t \leqslant \tau + \varepsilon.$$

由经典的 Gronwall 引理有

$$\|\Delta\bar{x}(t)\| \leqslant \left(\int_\tau^t \|\Delta_v f(\bar{x}(s), \bar{u}(s), s)\|\mathrm{d}s\right)\exp(\ell(t-\tau)) \leqslant K\varepsilon$$

对 $t \in [\tau, \tau+\varepsilon]$ 成立, 其中 $K = K(v)$ 不依赖 ε 和 τ.

剩下来, 在最后一个区间 $[\tau+\varepsilon, b]$ 上估计 $\Delta\bar{x}(t)$, 它满足方程

$$\Delta\dot{\bar{x}}(t) = f(\bar{x}(t) + \Delta\bar{x}(t), \bar{u}(t), t) - f(\bar{x}(t), \bar{u}(t), t), \quad \text{其中} \quad \|\Delta\bar{x}(\tau+\varepsilon)\| \leqslant K\varepsilon.$$

该方程的解在积分 (6.62) 的意义下理解. 因为

$$\begin{aligned}\|\Delta\bar{x}(t)\| &\leqslant \|\Delta\bar{x}(\tau+\varepsilon)\| + \int_{\tau+\varepsilon}^t \|f(\bar{x}(s) + \Delta\bar{x}(s), \bar{u}(s), s) - f(\bar{x}(s), \bar{u}(s), s)\|\mathrm{d}s \\ &\leqslant K\varepsilon + \ell \int_{\tau+\varepsilon}^t \|\Delta\bar{x}(s)\|\mathrm{d}s, \quad \tau+\varepsilon \leqslant t \leqslant b,\end{aligned}$$

再次利用 Gronwall 引理, 如果必要的话增大 K, 就可在整个区间 $[a, b]$ 上得到 $\|\Delta\bar{x}(t)\|$ 的估计. △

现在可以证明关于自由端点控制问题最大值原理的定理 6.37 了.

关于自由端点问题定理 6.37 的证明 设 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 为问题 (6.63) 的最优解, 且 $p(\cdot)$ 为具有边界/横截性条件 (6.70) 的伴随系统 (6.67) 相应的解. 要证明最大值条件 (6.66) 对几乎处处 $t \in [a, b]$ 成立. 如若不然, 那么存在正测度集 $T \subset [a, b]$, 使得

$$H(\bar{x}(t), p(t), \bar{u}(t), t) < \sup_{u \in U} H(\bar{x}(t), p(t), u, t), \quad t \in T.$$

在所作的假设下, 利用关于可测选择的标准结果可找到可测映射 $v: T \to U$, 满足

$$\Delta_v H(t) := H(\bar{x}(t), p(t), v(t), t) - H(\bar{x}(t), p(t), \bar{u}(t), t) > 0, \quad t \in T.$$

设 $T_0 \subset [a, b]$ 为函数 $H(t)$ 在区间 $[a, b]$ 上的 Lebesgue 正则 (或近似连续) 点集, 由经典的 Denjoy 定理, 它在 $[a, b]$ 上具有完全测度. 给定 $\tau \in T_0$ 和 $\varepsilon > 0$, 考虑最优控制的针形变分

$$u(t) := \begin{cases} v(t), & t \in T_\varepsilon := [\tau, \tau+\varepsilon) \cap T_0, \\ \bar{u}(t), & t \in [a, b] \setminus T_\varepsilon, \end{cases}$$

并且将引理 6.42 中关于费用泛函的增量公式应用到 $\bar{u}(\cdot)$ 和 $u(\cdot)$ 上. 由该公式有关系式

$$\Delta J[\bar{u}] = -\int_\tau^{\tau+\varepsilon} \Delta_v H(t)\mathrm{d}t + \eta_1 + \eta_2 + \eta_3,$$

其中 $\Delta_v H(t)$ 为 Hamilton-Pontryagin 函数的正增量, 并且余项 η_i, $i=1,2,3$ 是引理 6.42 中沿着相应于针形变分 $u(\cdot)$ 的轨道增量 $\Delta\bar{x}(\cdot)$ 来定义的. 由引理 6.43 的证明, 并适当修改以照顾到 T_ε 上可变扰动 $v(\cdot)$ (而不是 (6.71) 中的常数扰动), 则有 $\|\Delta\bar{x}(t)\|=O(\varepsilon)$, $t\in[a,b]$. 于是

$$\eta_1=o(\|\Delta\bar{x}(b)\|)=o(\varepsilon),\quad \eta_2=-\int_a^b o(\|\Delta\bar{x}(t)\|)\mathrm{d}t=o(\varepsilon),$$

$$\begin{aligned}\eta_3&\leqslant\int_\tau^{\tau+\varepsilon}\left|\left\langle\frac{\partial\Delta_v H(\bar{x}(t),p(t),\bar{u}(t),t)}{\partial x},\Delta\bar{x}(t)\right\rangle\right|\mathrm{d}t\\&\leqslant K\varepsilon\int_\tau^{\tau+\varepsilon}\left\|\frac{\partial\Delta_v H(\bar{x}(t),p(t),\bar{u}(t),t)}{\partial x}\right\|\mathrm{d}t=o(\varepsilon).\end{aligned}$$

由 $\tau\in T_0$ 的选择, 它是函数 $\Delta_v H(t)$ 的 Lebesgue 正则点, 和 Bochner 积分的构造有

$$\int_\tau^{\tau+\varepsilon}\Delta_v H(t)\mathrm{d}t=\varepsilon\big[H(\bar{x}(\tau),p(\tau),v(\tau),\tau)-H(\bar{x}(\tau),p(\tau),\bar{u}(\tau),\tau)\big]+o(\varepsilon).$$

因此得到表达式

$$\Delta J[\bar{u}]=-\varepsilon\big[H(\bar{x}(\tau),p(\tau),v(\tau),\tau)-H(\bar{x}(\tau),p(\tau),\bar{u}(\tau),\tau)\big]+o(\varepsilon).$$

对所有充分小的 $\varepsilon>0$, 沿着最优控制 $\bar{u}(\cdot)$ 的上述针形变分, 这蕴涵着 $\Delta J[\bar{u}]<0$. 这显然和 $\bar{u}(\cdot)$ 为问题 (6.63) 的最优控制相矛盾. 证毕. △

6.3.3 不等式约束问题的横截性条件

从 6.3.2 小节中可以看到, 自由端点最优控制问题最大值原理的解析证明对于端点约束 (6.64) 和/或 (6.65) 的情形并不成立. 事实上, 那个证明并不关心这些相应于针形控制变分轨道关于约束的可行性. 处理端点约束问题要求更为成熟的技巧, 它涉及系统 (6.61) 可达集的几何性质及其与费用泛函和端点约束的相互作用. 问题的关键在于证明存在由给定最优轨道的可行端点变分生成的凸集, 它不与最优性“禁止”的某些凸集相交, 于是可以利用凸集分离定理. 利用最优控制的多针形变分, 这是可以办到的. 由 $[a,b]$ 上时间的连续性, 后者是可以实现的. 实际上, 它反映了连续时间控制问题的隐含凸性.

在这一小节中只考虑涉及不等式类型 (6.64) 端点约束的最优控制问题. 具有等式约束 (6.65) 的控制问题有些不同 (更为复杂); 将在 6.3.4 小节中进行研究. 这里的主要目的是对由可微函数给出的不等式约束情形, 导出定理 6.37 中最大值原理关系中的横截性条件 (6.68). 如 6.3.1 小节讨论的那样, 在更一般的控制问题中和更少的限制假设下, 横截性条件可化为 (6.68) 中的形式或可类似地导出.

需要强调的是, 虽然是在 Banach 状态空间中研究最优控制问题, 但在系统的轨道上它们只涉及有限多个端点约束. 这里发展的方法有利于这种情形的考虑 (这在某种程度上和约束集的有限余维数性质相关联; 请比照推论 6.29, 推论 6.24 和注 6.25), 并且在费用和约束函数的可微算子下处理端点变分的有限维图像, 因此可利用有限维空间中的凸集分离定理.

在本小节余下的部分中考虑具有不等式端点约束 (6.64) 的最优控制问题 (6.63), 并且固定该问题的最优解 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$. 不失一般性, 假设 $\varphi_i(\bar{x}(b))=0$ 对所有 $i=0,\cdots,m$ 成立. 由证明易见 (对不等式约束这是很常见的), 如果 $\varphi_i(\bar{x}(b))<0$ 对某 $i\in\{1,\cdots,m\}$ 成立, 那么 $\lambda_i=0$, 即相应的函数 φ_i 可以从考虑之中排除. 此时定理 6.37 的互补松弛条件自动成立, 从而只需建立关系式 (6.66)~(6.68), 其中 $r=0$ 和 $0\neq(\lambda_0,\cdots,\lambda_m)\in\mathbb{R}^m$.

与 6.3.2 小节引入的 (单独) 针形变分相对应, 现在使用“多针形变分”, 其构造如下. 固定自然数 $M\geqslant 1$ 和初始时间区间中的 M 个点 $\tau_j\in[a,b]$, 使得 $a\leqslant\tau_1<\tau_2\leqslant\cdots<\tau_M<b$. 对 $j=1,\cdots,M$, 也考虑任意数 $N_j\in\mathbb{N}$ 和 $\alpha_{ij}\in[0,1]$, $i=1,\cdots,N_j$, 使得关系式

$$\tau_j+\varepsilon_0\sum_{i=1}^{N_j}\alpha_{ij}<\tau_{j+1},\quad j=1,\cdots,M-1,\quad \text{和}\quad \tau_M+\varepsilon_0\sum_{i=1}^{N_M}\alpha_{iM}<b$$

对某 $\varepsilon_0>0$ 成立. 下面构造参考控制 $\bar{u}(\cdot)$ 的扰动 $u(\cdot)$, 它在总长度较小的 $N_1+\cdots+N_M$ 个时间区间上与 $\bar{u}(\cdot)$ 不同, 而 $u(\cdot)$ 和 $\bar{u}(\cdot)$ 在这些区间上的差为可行控制区域 U 中的任意元素. 为此, 任取 $v_{ij}\in U$ 和 $\varepsilon\in(0,\varepsilon_0]$, 并且定义参考控制 $\bar{u}(\cdot)$ 的多针形变 $u(\cdot)$ 为

$$u(t):=\begin{cases} v_{ij}, & t\in\Big[\tau_j+\displaystyle\sum_{\nu=0}^{i-1}\alpha_{\nu j}\varepsilon,\tau_j+\sum_{\nu=1}^{i}\alpha_{\nu j}\varepsilon\Big),\quad \alpha_{0j}:=0,\ i=1,\cdots,N_j,\\ \bar{u}(t), & t\notin\Big[\tau_j,\tau_j+\displaystyle\sum_{i=1}^{N_j}\alpha_{ij}\varepsilon\Big),\quad j=1,\cdots,M. \end{cases} \tag{6.72}$$

注意到, 虽然有 M 个基本点 τ_j, 但是多针形变分 (6.72) 涉及 $N_1+\cdots+N_M$ 个针形类型扰动的点; 这与单针形变分 (6.71) 不同 (即使对 $M=1$ 的情形亦如此). 实际上, 多针形变分 (6.72) 是 $N_1+\cdots+N_M$ 个具有给定参数 $(\tau_j,v_{ij},\alpha_{ij},\varepsilon)$ 的单针形变分类型 (6.71) 的集合.

设 $\Delta\bar{x}_{\tau_j,v_{ij},\alpha_{ij},\varepsilon}(b)$ 为相应于单针形变分类型 (6.71) 并且具有参数 $(\tau_j,v_{ij},\alpha_{ij},\varepsilon)$ 的轨道 $\bar{x}(\cdot)$ 的端点增量. 像引理 6.43 的证明那样, 在处理光滑动态微分方程 (6.71) 及其沿着过程 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 关于 x 的线性化时, 可以验证 $\Delta\bar{x}_{\tau_j,v_{ij},\alpha_{ij},\varepsilon}(b)$ 和相应的

线性化端点增量 $\Lambda\bar{x}_{\tau_j,v_{ij},\alpha_{ij}}(b)$ 之间的关系式

$$\Delta\bar{x}_{\tau_j,v_{ij},\alpha_{ij},\varepsilon}(b) = [\alpha_{ij}\Lambda\bar{x}_{\tau_j,v_{ij},1}(b)]\varepsilon + o(\varepsilon), \tag{6.73}$$

其中 $\Lambda\bar{x}_{\tau_j,v_{ij},\alpha_{ij}}(b)$ 通过 (6.61) 沿着 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 关于 x 的线性化齐次方程

$$\dot{x} = \nabla_x f(\bar{x}(t),\bar{u}(t),t)x$$

的预解式 (Green 函数) $R(t,\tau)$ 计算为

$$\Lambda\bar{x}_{\tau_j,v_{ij},\alpha_{ij}}(b) = \alpha_{ij}R(b,\tau_j)\Delta_{v_{ij}}f(\bar{x}(\tau_j),\bar{u}(\tau_j),\tau_j) =: \alpha_{ij}\Lambda\bar{x}_{\tau_j,v_{ij},1}.$$

更进一步, 由多针形变分 (6.72) 生成的端点增量 $\Delta\bar{x}(b)$ 表示为

$$\Delta\bar{x}(b) = \Big[\sum_{j=1}^{M}\sum_{i=1}^{N_j}\alpha_{ij}\Lambda\bar{x}_{\tau_j,v_{ij},1}(b)\Big]\varepsilon + o(\varepsilon).$$

现在构造由涉及费用、约束函数导数以及相应参考最优控制 $\bar{u}(\cdot)$ 的所有多针形变分 (6.72) 的线性化端点增量的内积生成的有限维线性图像集合为

$$\begin{aligned} S := \Big\{(y_0,\cdots,y_m)\in\mathbb{R}^{m+1}\Big| \\ y_0 = \sum_{j=1}^{M}\sum_{i=1}^{N_j}\big\langle\nabla\varphi_0(\bar{x}(b)),\Lambda\bar{x}_{\tau_j,v_{ij},\alpha_{ij}}(b)\big\rangle,\cdots, \\ y_m = \sum_{j=1}^{M}\sum_{i=1}^{N_j}\big\langle\nabla\varphi_m(\bar{x}(b)),\Lambda\bar{x}_{\tau_j,v_{ij},\alpha_{ij}}(b)\big\rangle\Big\}, \end{aligned} \tag{6.74}$$

其中 $\tau_j\in[a,b)$, $v_{ij}\in U$, $\alpha_{ij}\in[0,1]$, $i=1,\cdots,N_j$, $N_j\in\mathbb{N}$, $j=1,\cdots,M$ 和 $M\in\mathbb{N}$.

关于 (6.74) 中的集合 S, 有两个关键的事实. 首先, 它是凸的, 这主要是因为在多针形变分 (6.72) 中利用任意 $\alpha_{ij}\in[0,1]$ 的可能性. 它是基于 $[a,b]$ 上关于时间的连续性和如上所述的反映了连续时间控制系统的内在凸性. 其次因为 $\bar{u}(\cdot)$ 在约束控制问题 (6.63) 和 (6.64) 中的最优性, 它确保了线性化图像集 (6.74) 和如下禁止点构成的凸集 (从问题的最优性和不等式约束的角度来看) 不相交

$$\mathbb{R}^{m+1}_{<} := \{(y_0,\cdots,y_m)\in\mathbb{R}^{m+1}\mid y_i<0,\quad\forall\, i=0,\cdots,m\}.$$

这两个事实将在下面的引理中得到证明.

引理 6.44 (不等式约束最优控制问题的隐含凸性和本原最优条件) 设 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 为不等式约束问题 (6.63) 和 (6.64) 的最优解. 除了 6.3.1 小节中的假设外, 还设所有的函数 φ_i 在点 $\bar{x}(b)$ 是 Fréchet可微的. 那么 (6.74) 中的线性化图像集 S 是凸的, 且和禁止点集 $\mathbb{R}^{m+1}_{<}$ 不相交.

证明 固定参数集 (τ_i, v_{ij}, N_j, M), 下面证明当数 α_{ij} 在 $[0,1]$ 中任取时, 集合 (6.74), 仍记为 S, 是凸的. 这显然蕴涵着"完全"集合 S 的凸性. 事实上, 选取由 (τ_i, v_{ij}, N_j, M) 组成的两个不同的集合, 总可将它们统一在一起得出一个容许多针形变分 (6.72). 于是只需对 α_{ij} 在区间 $[0,1]$ 上取值的情形证明 S 的凸性.

为此, 固定 (τ_i, v_{ij}, N_j, M) 以及两个集合 $\{\alpha_{ij}^{(1)}\}$ 和 $\{\alpha_{ij}^{(2)}\}$, 使得 (6.74) 中相应的点 $y^{(1)}$ 和 $y^{(2)}$ 属于线性化图像集 S. 那么对任意 $\lambda \in [0,1]$, 考虑点 $\lambda y^{(1)} + (1-\lambda) y^{(2)}$, 并且考虑到 $\Lambda \bar{x}_{\tau_i, v_{ij}, \alpha_{ij}}(b)$ 线性依赖于 α_{ij}, 可断言 $\lambda y^{(1)} + (1-\lambda) y^{(2)}$ 是 S 中的一个元素, 它相应于 $\{\lambda \alpha_{ij}^{(1)} + (1-\lambda) \alpha_{ij}^{(2)}\}$. 这就证明了 S 的凸性.

剩下来证明 $S \cap \mathbb{R}_{<}^{m_1} = \varnothing$, 其中 S 表示 (6.74) 中的"完全"图像集, 它相应于所有容许多针形变分 (6.72). 如若不然, 可找到一个具有某容许参数 $(\tau_i, v_{ij}, \alpha_{ij}, N_j, M)$ 的多针形变分 (6.72), 使得

$$\sum_{j=1}^{M} \sum_{i=1}^{N_j} \left\langle \nabla \varphi_0(\bar{x}(b)), \Lambda \bar{x}_{\tau_j, v_{ij}, \alpha_{ij}}(b) \right\rangle < 0, \cdots,$$

$$\sum_{j=1}^{M} \sum_{i=1}^{N_j} \left\langle \nabla \varphi_m(\bar{x}(b)), \Lambda \bar{x}_{\tau_j, v_{ij}, \alpha_{ij}}(b) \right\rangle < 0.$$

那么利用函数 $\varphi_0, \cdots, \varphi_m$ 在点 $\bar{x}(b)$ 的 Fréchet可微性、由 (6.72) 生成的端点增量 $\Delta \bar{x}(b)$, 以及对应于每个 $(\tau_j, v_{ij}, \alpha_{ij}, N_j, M)$ 的线性化 $\Lambda_{\tau_j, v_{ij}, \alpha_{ij}}$ 之间的上述关系, 对所有 $k = 0, \cdots, m$ 和所有充分小的 $\varepsilon > 0$, 有

$$\begin{aligned} \varphi_k(x(b)) - \varphi_k(\bar{x}(b)) &= \langle \nabla \varphi_k(\bar{x}(b)), \Delta \bar{x}(b) \rangle + o(\varepsilon) \\ &= \left[\sum_{j=1}^{M} \sum_{i=1}^{N_j} \left\langle \nabla \varphi_k(\bar{x}(b)), \Lambda \bar{x}_{\tau_j, v_{ij}, \alpha_{ij}}(b) \right\rangle \right] \varepsilon + o(\varepsilon) < 0. \end{aligned}$$

这意味着存在一个多针形控制变分 (6.72), 使得相应的轨道 $x(\cdot)$ 满足所有不等式约束 (6.64), 因此它是所考虑问题的可行轨道, 与 $\bar{x}(\cdot)$ 相比较, 轨道 $x(\cdot)$ 给出了 (6.63) 中费用泛函更小的值. 这与过程 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 在问题 (6.63) 和 (6.64) 中的最优性相矛盾. 证毕. △

所得到的关系式 $S \cap \mathbb{R}_{<}^{m+1} = \varnothing$ 可视为一个本原必要最优性条件, 当然并不是有效的, 因为它依赖于控制变分, 且不是由问题的初始数据来表述的. 为此, 利用凸集分离定理, 然后由引理 6.42 中增量方法的构造, 利用 Hamilton-Pontryagin 函数, 取它的对偶形式; 参见下面的讨论.

具有不等式约束问题定理 6.37 的证明 对引理 6.44 中的凸集 S 和 $\mathbb{R}_{<}^{m+1}$, 利用经典的分离定理, 可找到非零向量 $(\lambda_0, \cdots, \lambda_m) \in \mathbb{R}^{m+1}$, 使得

$$\sum_{i=0}^{m}\lambda_i y_i \geqslant \sum_{i=0}^{m}\lambda_i z_i, \quad \forall\ (y_0,\cdots,y_m)\in S \quad 和 \quad (z_0,\cdots,z_m)\in \mathbb{R}^{m+1}_{<}.$$

这明显蕴涵着对所有 $i=0,\cdots,m$, $\lambda_i \geqslant 0$, 且

$$\sum_{i=0}^{m}\lambda_i y_i \geqslant 0, \quad 只要 \quad (y_0,\cdots,y_m)\in S. \tag{6.75}$$

注意到, 向量 $(\lambda_0,\cdots,\lambda_m)$ 不依赖任何具体的多针形变分 (6.72); 它将所有这样的变分集合与 $0\in\mathbb{R}^{m+1}$ 分离开. 特别地, 对由具有参数 (τ,v,ε) 的单针形变分 (6.71) 生成的向量 $(y_0,\cdots,y_m)$. 利用 (6.75) 且考虑到沿着 (单) 针形变分最优轨道的完全和线性化增量之间的关系 (6.73), 对所有 $\tau\in[a,b)$, $v\in U$ 和充分小的 $\varepsilon>0$, 有

$$\sum_{i=0}^{m}\lambda_i\left\langle\nabla\varphi_i(\bar{x}(b)),\Delta_{\tau,v,\varepsilon}\bar{x}(b)\right\rangle+o(\varepsilon)\geqslant 0.$$

现在记

$$p(b):=-\sum_{i=0}^{m}\lambda_i\nabla\varphi_i(\bar{x}(b)),$$

并遵循引理 6.42 和 6.3.2 小节中关于自由端点控制问题定理 6.37 的证明, 将边界条件 (6.70) 替换为该条件, 即可完成证明. △

6.3.4 等式约束问题的横截性条件

为了完成定理 6.37 的证明, 剩下来证明问题中等式端点约束的情形. 不失一般性, 这里重点放在 (6.63) 和 (6.65) 给出的最优控制问题上, 即没有 6.3.3 小节中考虑过的不等式约束问题. 为了方便起见, 假设等式约束由前 m 个函数 φ_i 给出, 即

$$\varphi_i(x(b))=0, \quad i=1,\cdots,m. \tag{6.76}$$

接着, 再次构造 (6.74) 中的线性化映像集 S, 现在它是在费用和等式约束函数的梯度映射下由多针形变分的映像生成的. 在等式约束中, 禁止点的集合由

$$S^{<}:=\left\{(y_0,\cdots,y_m)\in\mathbb{R}^{m+1}\middle| y_0<0,y_1=0,\cdots,y_m=0\right\}$$

给出. 这里的目标是探寻映像集 S 和上面由 $\{\bar{u}(\cdot),\bar{x}(\cdot)\}$ 的最优性允许的禁止点集之间所有可能的关系. 最为困难的情形在下一个引理中考虑, 它建立了原点不可能是 S 的 $\mathbb{R}^m$- 投影的内点. 下面给出的证明涉及 Brouwer 不动点定理. 注意到, 虽然这个基本的拓扑结果本质是有限维的, 但是可以用它来处理无穷维空间中由发展型方程描述的最优控制问题. 正如所提及的那样, 关键在于该控制问题有有限多个端点约束, 这确保了约束集的有限余维数性质.

引理 6.45(等式约束下的端点变分)　在关于 X, U 和 f 的常驻假设下, 设 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 为控制问题 (6.63) 和 (6.76) 的最优解. 还假设函数 $\varphi_0, \cdots, \varphi_m$ 在点 $\bar{x}(b)$ 处 Fréchet可微, 且 $\varphi_1, \cdots, \varphi_m$ 在该点附近连续. 那么有

$$0 \notin \operatorname{int}(\operatorname{proj}_{\mathbb{R}^m} S),$$

其中线性化映像集合 S 由端点等式约束 (6.76) 在 (6.74) 中生成.

证明　假设结论不成立. 记 B_η 为 $\mathbb{R}^m$ 中半径为 $\eta > 0$, 中心在原点的闭球. 设 $\mathcal{T}$ 为内接于 B_η 的高维正“四面体”, 其顶点为 $q^{(s)}$, $s = 1, \cdots, m+1$. 如果 η 充分小, 那么对每个 $s = 1, \cdots, m+1$, 在多针形变分 (6.72) 中存在数 $\{\alpha_{ij}^{(s)}\}$ 和 $\nu < 0$, 使得对所有 $k = 1, \cdots, m$, 有

$$\begin{cases} \displaystyle\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle \nabla\varphi_0(\bar{x}(b)), \Lambda_{\tau_j, v_{ij}, \alpha_{ij}^{(s)}}\bar{x}(b)\right\rangle < \nu < 0, \\ \displaystyle q_k^{(s)} = \sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle \nabla\varphi_k(\bar{x}(b)), \Lambda_{\tau_j, v_{ij}, \alpha_{ij}^{(s)}}\bar{x}(b)\right\rangle, \end{cases}$$

其中 $q_k^{(s)}$ 表示顶点 $q^{(s)}$ 的第 k 个分量. 每个点 $q = q(\beta) \in \mathcal{T}$ 可表示为四面体顶点的凸组合

$$q(\gamma) = \sum_{s=1}^{m+1}\gamma_s q^{(s)} \quad 并且 \quad \gamma = (\gamma_1, \cdots, \gamma_{m+1}) \in P,$$

其中 P 表示 m- 维单纯形. 设 $u_{\gamma,\varepsilon}(\cdot)$ 为具体参数 $(\tau_j, v_{ij}, \alpha_{ij}(\gamma), \varepsilon)$ 的多针形变分 (6.72), 其中

$$\alpha_{ij}(\gamma) := \sum_{s=1}^{m+1}\gamma_s\alpha_{ij}^{(s)}, \quad \gamma = (\gamma_1, \cdots, \gamma_m) \in P.$$

现在考虑 ε- 参数映射族 $g(\cdot, \varepsilon) : P \to \mathbb{R}^m$, 定义为

$$g(\gamma, \varepsilon) := \left(\frac{\varphi_1(x_{\gamma,\varepsilon}(b)) - \varphi_1(\bar{x}(b))}{\varepsilon}, \cdots, \frac{\varphi_m(x_{\gamma,\varepsilon}(b)) - \varphi_m(\bar{x}(b))}{\varepsilon}\right),$$

其中 $x_{\gamma,\varepsilon}(\cdot)$ 表示 (6.61) 的相应于多针形控制变分 $u_{\gamma,\varepsilon}(\cdot)$ 的轨道. 再记

$$\begin{aligned} g(\gamma, 0) := \Bigg(&\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle \nabla\varphi_1(\bar{x}(b)), \alpha_{ij}(\gamma)\Lambda\bar{x}_{\tau_j, v_{ij}, 1}(b)\right\rangle, \cdots, \\ &\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle \nabla\varphi_m(\bar{x}(b)), \alpha_{ij}(\gamma)\Lambda\bar{x}_{\tau_j, v_{ij}, 1}(b)\right\rangle\Bigg), \end{aligned}$$

可断言, 映射 $g(\cdot,\cdot)$ 在 $P\times[0,\varepsilon_0]$ 上连续, 其中 ε_0 充分小. 这是根据 $\varphi_1,\cdots,\varphi_m$ 在点 $\bar{x}(b)$ 的 Fréchet可微性和这些函数在该点附近的连续性. 由上面的构造可得

$$g(\gamma,0)=\sum_{s=1}^{m+1}\gamma_s q^{(s)},\quad g(P,0)=\mathcal{T};$$

因此, 集合 $g(P,0)$ 包含原点作为一个内点. 现在证明存在 $\widehat{\varepsilon}>0$, 使得

$$0\in \mathrm{int}g(P,\varepsilon)\quad (\forall\varepsilon<\widehat{\varepsilon}).$$

为此, 注意到映射 $g(\cdot,0)$ 是从 P 到 $\mathcal{T}$ 中的 1-1 和连续映射. 于是它的逆映射是单值和连续的; 记为 $p(y)$, 且令

$$h(y,\varepsilon):=g(p(y),\varepsilon)\quad (\forall y\in\mathcal{T}\quad \text{和}\quad \varepsilon\in[0,\varepsilon_0]).$$

取 $\eta>0$ 充分小使得半径为 η, 中心在原点的球 B_η 属于四面体 $\mathcal{T}$. 那么由映射 $h(\cdot,\cdot)$ 的连续性可得, 存在 $\widehat{\varepsilon}>0$, 使得

$$\|h(y,0)-h(y,\varepsilon)\|<\eta\quad \text{只要}\quad \varepsilon<\widehat{\varepsilon}.$$

因此, 给定任意 $\varepsilon\in(0,\widehat{\varepsilon})$, 连续映射 $h(y,0)-h(y,\varepsilon)$ 将球 B_η 映为它本身. 利用 Brouwer 不动点定理, 可找到点 $y^\varepsilon\in B_\eta$ 满足

$$h(y^\varepsilon,0)-h(y^\varepsilon,\varepsilon)=y^\varepsilon\quad (\forall\varepsilon\in(0,\widehat{\varepsilon})).$$

由 $h(y,0)\equiv y$, 这蕴涵着

$$h(y^\varepsilon,\varepsilon)=g(p(y^\varepsilon),\varepsilon)=g(\gamma^\varepsilon,\varepsilon)$$

对某满足 $g(\gamma^\varepsilon,0)=y^\varepsilon$ 的 $\gamma^\varepsilon\in P$ 成立. 考虑到 $g(\cdot,\cdot)$ 的构造, 可断言, 对所有 $\varepsilon\in(0,\widehat{\varepsilon})$, 由多针形变分 $u_{\gamma^\varepsilon,\varepsilon}(\cdot)$ 生成的轨道 $x_{\gamma^\varepsilon,\varepsilon}(\cdot)$ 满足等式约束 (6.76). 此外, 对于沿着费用泛函的变分有

$$\begin{aligned}&\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle\nabla\varphi_0(\bar{x}(b)),\varLambda_{\tau_j,v_{ij},\alpha_{ij}(\gamma^\varepsilon)}\bar{x}(b)\right\rangle\\ =&\sum_{s=1}^{m+1}\gamma_s^\varepsilon\left(\sum_{j=1}^{M}\sum_{i=1}^{N_j}\left\langle\nabla\varphi_0(\bar{x}(b)),\varLambda_{\tau_j,v_{ij},\alpha_{ij}^{(s)}}\bar{x}(b)\right\rangle\right)\\ <&\sum_{s=1}^{m+1}\gamma_s^\varepsilon\nu<\nu\quad(\forall\varepsilon\in(0,\widehat{\varepsilon})).\end{aligned}$$

类似不等式约束的情形, 沿着等式约束问题 (6.63) 和 (6.65) 的某可行解, 上式蕴涵着

$$\varphi_0(x_{\gamma^\varepsilon,\varepsilon}(b)) < \varphi_0(\bar{x}(b)).$$

这和问题中过程 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 的最优性相矛盾. 证毕. △

根据引理 6.45 以及 6.3.2 和 6.3.3 小节中的讨论, 最后证明定理 6.37 中剩下的等式约束的情形. 这样就完成了整个定理的证明.

关于等式约束问题定理 6.37 的证明　考虑到引理 6.45 和 (6.72) 中相应于等式约束 (6.76) 的线性化映像集 S 和禁止点的集合 $S^<$ 之间有如下两种可能的关系:

(a) $S \cap S^< = \varnothing$;

(b) $S \cap S^< \neq \varnothing$ 和 $0 \in \mathrm{bd}(\mathrm{proj}_{\mathbb{R}^m} S)$.

首先考虑情形 (a). 因为集合 S 和 $S^<$ 都是凸的, 利用经典的凸集分离定理, 可找到一个非零向量 $(\lambda_0, \cdots, \lambda_m) \in \mathbb{R}^{m+1}$, 使得

$$\sum_{i=0}^{m} \lambda_i y_i \geqslant \sum_{i=0}^{m} \lambda_i z_i$$

对所有 $(y_0, \cdots, y_m) \in S$ 和 $(z_0, \cdots, z_m) \in S^<$ 成立. 由禁止集合 $S^<$ 的结构, 这明显蕴涵着 $\lambda_0 \geqslant 0$ 和关系式 (6.75). 现在像 6.3.3 小节末对不等式约束情形那样作同样处理, 则完成该情形定理的证明.

剩下验证情形 (b). 记 $\Omega := \mathrm{proj}_{\mathbb{R}^m} S$ 并注意到在 $\mathbb{R}^m$ 中该集合是闭凸的. 因为 $0 \in \mathrm{bd}\Omega$, 利用关于凸集的支撑超平面定理可找到一个非零 m- 向量 $(\lambda_1, \cdots, \lambda_m)$ 在原点支撑 Ω. 那么又可得到具有非平凡 $(m+1)$- 向量 $(0, \lambda_1, \cdots, \lambda_m)$ 的基本关系式 (6.75). 类似不等式约束的情形, 就完成了定理的证明. △

注意到, 等式约束函数 φ_i 除了在 $\bar{x}(b)$ 点的 Fréchet可微性条件外, 它们在该点附近的连续性假设对定理 6.37 也是至关重要的, 甚至对有限维状态空间 X 且具有平凡动态 $f = 0$ 的情形亦如此; 参见例 5.12.

6.4　最优控制中的近似最大值原理

本节研究由连续时间控制系统的离散逼近控制的参数族动态系统的最优控制问题. 离散/有限差分逼近在最优控制的定性和数值方面都是至关重要的. 当离散/有限差分逼近被视为一个缩减步骤的离散化过程时, 它们处于连续时间控制系统和固定步骤的离散时间控制系统的中间位置. 微分包含离散逼近的一般控制问题在 6.1 节中已研究过, 但那里的观念和这一节中的不一样, 即从离散到连续: 建立具有固定离散化步骤的离散时间系统的必要最优性条件, 然后利用适定的离散逼近来

导出连续时间控制系统的最优性条件. 6.1 节在某些关于系统动态的凸性/松弛的先验假设下, 得到的结果提供了最大值原理类型的必要条件.

现在要在离散时间和连续时间控制系统的关系上探索相反的方向, 即从连续到离散, 在没有任何凸性/松弛假设的连续时间动态上, Pontryagin 最大值原理 (PMP) 及其在非光滑问题和微分包含上的扩展是成立的. 但要阐明建立离散逼近最大值原理类型的必要最优性条件的可能性是极富挑战的. 在这个方向上得到的结果是相当出人意外的, 结果如下.

6.4.1 离散时间控制系统的确切和近似最大值原理

正如在 6.2 和 6.3 节中看到的那样, Weierstrass-Pontryagin 最大值条件的最大值原理, 对没有先验凸性假设的连续时间控制系统是成立的. 这是由于连续时间动态的具体特性产生了在这样的控制系统中固有的隐含凸性. 或许最为显著和深刻阐述连续时间系统隐含凸性的结果是关于无原子/连续向量测度值域的 Lyapunov 定理. 它等价于集值积分的 Aumann 凸性定理; 参见引理 6.18 证明中的讨论及其中的参考文献. 在 6.3 节给出的具体光滑动态控制系统的最大值原理的证明中没有用到这些结果, 而是在构造针形 (和多针形) 变分中直接利用了时间的连续性, 从而导致了线性化像集的自动凸性, 如引理 6.44. 但对由一般离散包含

$$x(t+1) \in F(x(t),t), \quad t=0,\cdots,K-1,$$

或由它们的参数化控制表示

$$x(t+1) = f(x(t),u(t),t), \quad u(t) \in U, \quad t=0,\cdots,K-1,$$

描述的离散时间系统则不能期望得到这样的性质, 其中 $K \in \mathbb{N}$ 表示离散动态过程步骤的次数 (最终的离散时间). 然而, 如果"离散速率"集合 $F(x,t)$ 或参数化控制系统中的 $f(x,U,t)$ 是凸的, 那么离散最大值原理成立. 在此情形中, 最大值条件实际上是上面讨论过的 Euler-Lagrange 包含的一个直接推论. 事实上, 由于凸集法锥的特殊表示, 这可由定理 1.34 中凸值映射上导数的极值性质得到.

众所周知, 如果上面的速率集合不是凸的, 那么离散最大值原理可能不成立, 甚至对简单的光滑动态控制系统亦如此. 现在对一族具有光滑动态的简单自由端点问题, 给出一个离散最大值原理 (Pontragin 最大值原理在离散时间控制系统上的一个自然类比) 不成立的例子. 在这个例子中, Hamilton-Pontryagin 函数沿着任意可行控制达到整体最小值 (而不是最大值). 与本章其他地方一样, 这里"自由端点"问题的意思是在系统轨道的右端点没有约束, 但左端点可能固定.

例 6.46(离散最大值原理不成立) 在具有光滑动态和无端点约束的二维离散控制系统上, 存在极小化线性函数的一族最优控制问题使得这些问题的任意可行控制不满足离散最大值原理.

证明 考虑下面具有二维状态向量 $x=(x_1,x_2)\in\mathbb{R}^2$ 的最优控制问题族:

$$\begin{cases}\min & J[u,x]=\varphi(x(K)):=x_2(3)\\ \text{s.t.} & x_1(t+1)=\vartheta(u(t),t),\quad x_1(0)=0,\\ & x_2(t+1)=\gamma(x_1(t))^2+\eta x_2(t)-(\gamma/\eta)(\vartheta(u(t),t))^2,\quad x_2(0)=0,\\ & u(t)\in U,\quad t=0,1,2,\end{cases}$$

其中标量函数 $\vartheta(\cdot,\cdot)$, 数 γ, η 和控制集 U 是任意的, 那么该系统 (一个自然离散) 的 Hamilton-Pontryagin 函数为

$$\begin{aligned}H(x(t),p(t+1),u,t):=&\langle p(t+1),f(x(t),u,t)\rangle\\ =&p_1(t+1)\vartheta(u,t)+\gamma p_2(t+1)(x_1(t))^2\\ &+\eta p_2(t+1)x_2(t)-(\gamma/\eta)p_2(t+1)(\vartheta(u,t))^2,\end{aligned}$$

其中伴随轨道 $p(\cdot)$ 满足系统 (6.67) 相应的离散化

$$p(t)=\nabla_x H(x(t),p(t+1),u(t),t),\quad t\in\{0,\cdots,K-1\}=\{0,1,2\},$$

具有边界/横截性条件

$$p(K)=-\nabla\varphi(x(K))=(0,-1)\quad(K=3).$$

对所考虑的问题有

$$\begin{aligned}&p_2(3)=-1,\quad p_2(2)=-\eta,\quad p_2(1)=-\eta^2,\\ &p_1(3)=0,\quad p_1(2)=-\gamma x_1(2)=-2\gamma\vartheta(u(1),1),\\ &p_1(1)=-2\gamma\eta x_1(1)=-2\gamma\eta\vartheta(u(0),0).\end{aligned}$$

那么在 Hamilton-Pontryagin 函数中, 只考虑依赖于 u 的项, 可得到

$$\begin{aligned}&H(u,0)=-\gamma\eta[2\vartheta(u(0),0)\vartheta(u,0)-(\vartheta(u,0))^2],\\ &H(u,1)=-\gamma[2\vartheta(u(1),1)\vartheta(u,1)-(\vartheta(u,1))^2].\end{aligned}$$

这表明, 给定任意 $\vartheta(\cdot,\cdot)$ 和 U, 只要 $\gamma>0$ 和 $\gamma\eta>0$, 函数 $H(u,0)$ 和 $H(u,1)$ 分别在任意的 $u(0)$ 和 $u(1)$ 达到整体最小值. 因此在所考虑的最优控制问题族中, 上面的离散最大值原理的关系对最优性不是必要的. 证毕. △

值得一提的是, 当 $t=K-1=2$ 时, 例 6.46 中的 Hamilton-Pontryagin 函数对于最优控制在 $u\in U$ 上的确达到整体最大值. 这可将增量公式沿着最优控制的针形变分应用到凹费用函数上得到; 参阅 6.4.2 小节中的讨论. 此外, 离散最大值原理

在例 6.46 的问题族中对所有 t 都成立, 即它提供了沿着最优控制在每个时刻的必要最优性条件, 当且仅当

$$\gamma \leqslant 0 \quad 和 \quad \eta \geqslant 0.$$

这可由上面的讨论和 Mordukhovich 的书 [901] 中第 17 节中的结果得到, 其中给出了离散最大值原理的一些单独条件. 因此同时满足条件 $\gamma \leqslant 0$ 和 $\eta \geqslant 0$ 完全描述了例 6.46 中问题的初始数据, 它确保满足最大值原理. 注意到, 在前面提到的书 [901] 中, 在该方向得到的结果总体上看, 都大量利用了非凸离散时间控制系统的初始数据之间的相互联系; 参见其中更多的讨论和例子.

本节的注意力主要不是放在由具有固定离散时间动态系统控制的最优控制问题上, 而是放在 6.3 节中研究过的连续时间问题的有限差分/离散逼近上. 这意味着不是考虑连续时间控制系统 (6.61), 而是考虑它的一系列差分

$$\begin{cases} x_N(t+h_N) = x_N(t) + h_N f(x_N(t), u_N(t), t), \quad x_N(a) = x_0 \in X, \\ u_N(t) \in U, \quad t \in T_N := \{a, a+h_N, \cdots, b-h_N\}, \end{cases} \tag{6.77}$$

其中 $N \in \mathbb{N}$ 和 $h_N := (b-a)/N$. 微分/发展包含的离散逼近在 6.1 节中已研究过, 并用来作为导出连续时间控制问题必要最优条件的工具. 现在的目标截然相反: 从连续问题的角度来看离散逼近的最优控制问题. 关键的问题是:

当 $N \to \infty$ 时, 是否可能得到由 (6.77) 类型的非凸有限差分系统控制的最优控制问题的 Pontryagin 最大值原理?

如果答案是否定的, 那么 PMP 的这种潜在的不稳定性可能会给它在涉及时间导数有限差分逼近的任何数值逼近的应用上带来严峻的挑战.

首先, 对每个 $N \in \mathbb{N}$, 考虑在满足 (6.77) 的离散时间过程 $\{u_N(\cdot), x_N(\cdot)\}$ 上, 极小化端点函数 $\varphi_0(x(b))$ 的问题. 对这样的每个问题的确切 PMP, 其相应的离散最大值原理可写为: 给定最优过程 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 存在伴随轨线 $p_N(\cdot)$, $t \in T_N \cup \{b\}$, 满足

$$p_N(t) = p_N(t+h_N) + h_N \nabla_x H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) \tag{6.78}$$

(当 $t \in T_N$ 时) 和横截性条件

$$p_N(b) = -\nabla \varphi_0(\bar{x}_N(b)), \tag{6.79}$$

使得当 $N \in \mathbb{N}$ 时, 确切最大值条件

$$H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) = \max_{u \in U} H(\bar{x}_N(t), p_N(t+h_N), u, t), \quad t \in T_N$$

成立, 其中 Hamilton-Pontryagin 函数为

$$H(x, p, u, t) := \langle p, f(x, u, t) \rangle.$$

由例 6.46 (通过标准的尺度调整) 和上面的讨论可知, 这个 (确切) 离散最大值原理一般不成立, 即使对由离散逼近系统类型 (6.77) 控制的简单最优控制问题亦可能对任意 $N \in \mathbb{N}$ 不成立. 这似乎表明在离散逼近下 PMP 的一种可能的不稳定性. 但事实上, 对连续时间控制系统离散化下 PMP 的稳定性, 若需要 PMP 的这种离散化确切版本成立, 则要求太过了.

这里真正需要的是沿着离散逼近问题的每一个最优解序列 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 当 $N \in \mathbb{N}$ 足够大时, 近似最大值条件, 即

$$H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) = \max_{u \in U} H(\bar{x}_N(t), p_N(t+h_N), u, t) + \varepsilon(t, h_N)$$

对所有 $t \in T_N$ 和某 $\varepsilon_N(t, h_N) \to 0$ (当 $N \to \infty$) 成立, 且对 $t \in T_N$ 是一致的, 其中 $p_N(\cdot)$ 是满足 (6.78) 和 (6.79) 相应的伴随轨道. 此时称所考虑的离散逼近问题的“近似最大值原理 (AMP)”成立. 这样的 PMP 近似版本确保了 PMP 的离散化稳定性, 因此表明了非凸连续时间控制系统在计算机计算和模拟中利用 PMP 的可能性. 更进一步, AMP 对离散逼近问题序列给出了必要最优性条件, 在求解具有充分小步长的离散时间控制问题时起到了和 (确切) 离散最大值原理本质上相同的作用; 参见例 6.68. 然而, 在大步长 h 的情形, 虽然近似最大值原理对最优性仍然是必要的, 但和确切最大值原理可能相去甚远.

在 6.4.3 小节中证明了, 在任何 Banach 状态空间 X 中, 对光滑自由端点最优控制问题, 即对在具有光滑动态和无端点约束的离散逼近系统 (6.77) 上极小化光滑 (连续可微) 费用函数的问题, AMP 在 $\varepsilon(h_N, t) = O(h_N)$ 时成立. 其证明是纯分析的, 其中利用了 (单) 针形控制变分和 6.3.2 小节中相应的离散增量公式.

连续时间系统的 PMP 和离散逼近 AMP 的重要区别在于, 即使对涉及最简单非光滑 (甚至凸) 费用函数的最优控制问题, AMP 竟然也没有预期的 (下) 次微分版本. 相应的反例在 6.4.3 小节中给出, 它与那些表明具有 Fréchet可微 (但不是连续可微) 费用函数的最优控制问题以及具有非光滑动态控制问题的 AMP 失效的反例一道给出.

因此, AMP 对非光滑性非常灵敏. 另一方面, 在 6.4.3 小节中导出了平行于 6.3.1 小节中连续时间系统 PMP 的上次微分版本的 AMP. 与连续时间系统的 PMP 相比较, 它对具有更多限制的费用函数类成立. 这类一致上次可微函数将在 6.4.2 小节中引入和研究.

具有端点约束离散逼近系统的最优控制问题要复杂得多. 考虑具有光滑不等式约束

$$\varphi_i(x_N(b)) \leqslant 0, \quad i = 1, \cdots, m$$

的控制系统, 在 6.4.4 小节中, 在对最优控制序列所作的一些恰当性假设下, 将描述具有扰动互补松弛条件的 AMP, 这可作为具有分段连续性的离散 AMP. 这里提

到的恰当性假设对非凸约束系统的 AMP 是本质的, 这将由例子表明. 6.4.5 小节中给出的 AMP 证明揭示了所考虑的有限差分控制问题的隐含凸性性质的一个近似版本; 参见如下更多的细节和讨论. 另外也将导出具有一致上次可微端点函数 φ_i $(i=0,\cdots,m)$ 的不等式约束问题 AMP 的上次微分形式.

具有等式约束

$$\varphi_i(x(b))=0,\quad i=m+1,\cdots,m+r$$

的连续时间控制问题离散逼近的适当设置涉及约束扰动

$$|\varphi_i(x_N(b))|\leqslant\xi_{iN},\quad i=m+1,\cdots,m+r,$$

其中当 $N\to\infty$ 时, $\xi_{iN}\downarrow 0$. 在 6.4.5 小节中证明了, 如果有下面的相容性条件 (consistency condition),

$$\limsup_{N\to\infty}\frac{h_N}{\xi_N}=0\quad(\forall i=m+1,\cdots,m+r),\tag{6.80}$$

那么由光滑函数描述的具有扰动等式约束离散逼近问题的 AMP 成立. 这意味着等式约束扰动 ξ_{iN} 比离散步长 h_N 趋于零要慢, 这特别要求 $\xi_{iN}\neq 0$. 这里给出了一个例子来证明相容性条件 (6.80) 对 AMP 是本质的, 即使当 $\xi_{iN}=O(h_N)$ 时, AMP 也可能不成立.

所得的结果可以扩展到具有时滞状态变量的离散逼近系统中, 这关系到时间长度 $b-a$ 和近似步长 h_N 的不可公度的情形; 参见 6.4.6 小节. 另一方面, 这里通过例子表明中立型系统离散逼近的 AMP 不成立, 即使在光滑自由端点控制问题的情形亦如此.

在导出所提到的关于 AMP 的结果之前, 先描述和研究 Banach 空间中一致上次可微函数类, 这将用来导出上次微分形式的 AMP. 这类函数特别地包含了连续可微函数以及凹连续函数, 它们在应用中是有特别意义的.

6.4.2 一致上次可微函数

本小节的主要研究对象是如下定义的一类函数.

定义 6.47(一致上次可微性) 称在 Banach 空间 X 上的实值函数在点 $\bar{x}$ 附近是"一致上次可微"的, 如果对 $\bar{x}$ 的某邻域 V 中的每个 x, 存在非空集合 $\mathcal{D}^+\varphi(x)\subset X^*$, 其描述为: 对任意给定的 $\varepsilon>0$, 存在 $\nu>0$, 使得 $x^*\in\mathcal{D}^+\varphi(x)$ 当且仅当

$$\varphi(v)-\varphi(x)-\langle x^*,v-x\rangle\leqslant\varepsilon\|v-x\|\tag{6.81}$$

对任意满足 $\|v-x\|\leqslant\nu$ 的 $v\in V$ 成立.

易见, 在任何 Banach 空间中, 这类函数包括了所有光滑 (即在 $\bar{x}$ 附近是 $\mathcal{C}^1$ 的) 函数 (此时 $\mathcal{D}^+\varphi(x) = \{\nabla\varphi(x)\}$), 同时也包括了所有凹连续函数 (此时对 $\bar{x}$ 附近的 x 有 $\mathcal{D}^+\varphi(x) = \partial^+\varphi(x)$). 更进一步, 由定义可导出上面的函数类关于在紧集上取极小值运算是闭的. 注意到, 即使 φ 在 $\bar{x}$ 附近是 Lipschitz 的并且在 $\bar{x}$ 是 Fréchet可微的, 它在该点附近也可能不是一致上次可微的. 一个简单的例子由标准函数 $\varphi : \mathbb{R} \to \mathbb{R}$ 给出, 它定义为 $\varphi(x) := x^2 \sin(1/x)$ (若 $x \neq 0$) 并且 $\varphi(0) := 0$ (若 $x = 0$).

在阐述本小节主要结果之前, 先考虑一个在 $\bar{x}$ 有限的任意函数 $\varphi : X \to \overline{\mathbb{R}}$, 并且描述 (1.52) 中定义的 φ 在 $\bar{x}$ 的 Fréchet上次微分

$$\widehat{\partial}^+\varphi(\bar{x}) := \left\{ x^* \in X^* \,\middle|\, \limsup_{x \to \bar{x}} \frac{\varphi(x) - \varphi(\bar{x}) - \langle x^*, x - \bar{x} \rangle}{\|x - \bar{x}\|} \leqslant 0 \right\}$$

和 φ 在 $\bar{x}$ 的 Dini (或 Dini-Hadamard) 上方向导数的两种形式的关系, 其中标准 (强) 形式定义为

$$d^+\varphi(\bar{x}; z) := \limsup_{\substack{y \to z \\ t \downarrow 0}} \frac{\varphi(\bar{x} + ty) - \varphi(\bar{x})}{t},$$

弱形式定义为

$$d_w^+\varphi(\bar{x}; z) := \limsup_{\substack{y \xrightarrow{w} z \\ t \downarrow 0}} \frac{\varphi(\bar{x} + ty) - \varphi(\bar{x})}{t},$$

其中 $y \xrightarrow{w} z$ 表示 X 中的弱收敛. 下面用到的命题就其自身而言当然也是有意义的, 它揭示了次梯度和方向导数构造之间的对偶性. 一般地, 这在自反空间中对弱方向导数成立并且在有限维空间中对强方向导数成立. 这里只描述本节中需要的上方构造; 它蕴涵着下方情形的对应结果.

命题 6.48(Fréchet次梯度和 Dini 方向导数的关系)　总有

$$\begin{aligned} \widehat{\partial}^+\varphi(\bar{x}) &\subset \{x^* \in X \mid \langle x^*, z \rangle \geqslant d_w^+\varphi(\bar{x}; z), \quad \forall\, z \in X\} \\ &\subset \{x^* \in X \mid \langle x^*, z \rangle \geqslant d^+\varphi(\bar{x}; z), \quad \forall\, z \in X\}, \end{aligned}$$

其中当 X 是自反时, 在第一个包含中等式成立, 而当 $\dim X < \infty$ 时, 在第二个包含中等式成立. 此外, 如果 φ 在 $\bar{x}$ 附近是局部 Lipschitz 的, 那么有

$$d^+\varphi(\bar{x}; z) := \limsup_{t \downarrow 0} \frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t}. \tag{6.82}$$

证明　为了证明命题中的最后一个包含, 只需注意到对每个 $x^* \in \widehat{\partial}^+\varphi(\bar{x})$ 和 $z \in X$ 有

$$d^+\varphi(\bar{x}; z) - \langle x^*, z \rangle = \limsup_{\substack{y \to z \\ t \downarrow 0}} \frac{\varphi(\bar{x} + ty) - \varphi(\bar{x}) - t\,\langle x^*, y \rangle}{t} \leqslant 0;$$

另外一个包含类似地证明. 现在证明如果 X 是自反的, 那么在第一个包含中等式成立. 为此, 选取 $x^* \notin \widehat{\partial}^+\varphi(\bar{x})$, 并且任取 $\gamma > 0$, 那么存在一个序列 $x_k \to \bar{x}$ 使得

$$\varphi(x_k) - \varphi(\bar{x}) - \langle x^*, x_k - \bar{x}\rangle - \gamma\|x_k - \bar{x}\| > 0, \quad \forall\, k \in \mathbb{N}.$$

因为 X 是自反的, 不失一般性, 假设序列 $(x_k - \bar{x})/\|x_k - \bar{x}\|$ 弱收敛于某 $z \in X$. 那么

$$d^+\varphi(\bar{x}; z) \geqslant \limsup_{k\to\infty} \frac{\varphi(x_k) - \varphi(\bar{x})}{\|x_k - \bar{x}\|} \geqslant \langle x^*, z\rangle + \gamma,$$

这就确保了所要的等式, 因为 γ 是任意选取的.

剩下来证明表示 (6.82), 如果 φ 在 $\bar{x}$ 附近是局部 Lipschitz 的且具有模 $\ell > 0$. 那么当 $t > 0$ 充分小时, 可得到

$$|\varphi(\bar{x} + ty) - \varphi(\bar{x} + tz)| \leqslant t\ell\|y - z\|, \quad \forall\, y, z \in X.$$

因此有

$$\begin{aligned} d^+\varphi(\bar{x}, z) &= \limsup_{\substack{y\to z\\ t\downarrow 0}} \left[\frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t} + \frac{\varphi(\bar{x} + ty) - \varphi(\bar{x} + tz)}{t}\right] \\ &= \limsup_{t\downarrow 0} \frac{\varphi(\bar{x} + tz) - \varphi(\bar{x})}{t} \quad (\forall z \in X), \end{aligned}$$

这就证明了 (6.82). 证毕. △

现在可以建立一致上次可微函数的重要性质了, 它们在下面要用到, 就其自身而言当然也是有意义的. 特别地, 它表明了这样的函数具有紧接着定义 1.91 后讨论的上正则性性质.

定理 6.49(一致上次可微函数的性质) 设 X 为自反的, 且设 φ 在 $\bar{x}$ 连续和在该点附近一致上次可微, 且由定义 6.47, 具有次梯度集 $\mathcal{D}^+\varphi(x)$. 那么存在 $\bar{x}$ 的一个邻域使得 φ 在该邻域上是 Lipschitz 的并且可选取

$$\mathcal{D}^+\varphi(x) = \widehat{\partial}^+\varphi(x) = \partial^+\varphi(x).$$

证明 梯度集 $\mathcal{D}^+\varphi(x)$ 显然是凸的. 此外, 容易验证这些集合中的每一个在 X^* 中都是范数闭的. 由它的凸性和 X 的自反性假设, 它也是弱闭的. 接下来证明 $\mathcal{D}^+\varphi(x)$ 在 X^* 中在 $\bar{x}$ 附近是一致有界的. 如若不然, 那么可选取序列 $x_k \to \bar{x}$ 和 $x_k^* \in \mathcal{D}^+\varphi(x)$, 使得 $\|x_k^*\| \to \infty$ $(k \to \infty)$. 那么利用 Hahn-Banach 定理并考虑到 X 的自反性, 可找到 $u_k \in X$ 满足关系式

$$\langle x_k^*, u_k\rangle = \|x_k^*\|^{1/2} \quad 和 \quad \|u_k\| = \|x_k^*\|^{-1/2}, \quad \forall\, k \in \mathbb{N}.$$

现在记 $v_k := x_k - u_k$, 由 (6.81) 有

$$\varphi(v_k) - \varphi(x_k) \leqslant -\langle x_k^*, u_k\rangle + \varepsilon\|u_k\|,$$

并且由上面的构造有 $\|u_k\| \to 0$ 和 $\langle x_k^*, u_k\rangle \to \infty$. 这就得到 $\varphi(v_k) - \varphi(x_k) \to -\infty$ 而且 x_k, $v_k \to \bar{x}$ $(k \to \infty)$, 这与 φ 在 $\bar{x}$ 的连续性相矛盾. 因此证明了 $\mathcal{D}^+\varphi(x)$ 在该点附近的一致有界性.

下面证明 φ 在 $\bar{x}$ 附近是局部 Lipschitz 的. 根据定理 3.49 的中值不等式, 这对 $\mathcal{D}^+\varphi(\cdot)$ 成立, 类似定理 3.52 的证明可得到. 然而, 直接利用集合 $\mathcal{D}^+\varphi(x)$ 在 $\bar{x}$ 附近的一致有界性和性质 (6.81) 可容易地给出证明. 事实上, 如若不然, 那么可找到序列 $x_k \to \bar{x}$ 和 $v_k \to \bar{x}$ 满足

$$|\varphi(v_k) - \varphi(x_k)| > k\|v_k - x_k\| \quad (k \to \infty).$$

明确起见, 假设 $\varphi(v_k) - \varphi(x_k) > k\|v_k - x_k\|$; 另外一种情形是对称的. 现在利用 φ 的一致上次可微性, 可找到序列 $x_k^* \in \mathcal{D}^+\varphi(x_k)$, 对任意给定的 $\varepsilon > 0$, 当 k 充分大时, 满足

$$\begin{aligned} k\|v_k - x_k\| &< \varphi(v_k) - \varphi(x_k) \leqslant \langle x_k^*, v_k - x_k\rangle + \varepsilon\|v_k - x_k\| \\ &\leqslant (\|x_k^*\| + \varepsilon)\|v_k - x_k\|. \end{aligned}$$

这就得到 $\|x_k^*\| \to \infty$ $(k \to \infty)$, 这与集合 $\mathcal{D}^+\varphi(x)$ 在 $\bar{x}$ 附近的一致有界性相矛盾. 因此证明了 φ 的局部 Lipschitz 性质.

由 (1.52) 中的 Fréchet上次梯度的定义和 (6.81) 中 $\mathcal{D}^+\varphi(x)$ 的构造有 $\mathcal{D}^+\varphi(x) \subset \widehat{\partial}^+\varphi(x)$. 下面证明 $\mathcal{D}^+\varphi(x) = \widehat{\partial}^+\varphi(x)$. 首先注意到, 在 $X \times X^*$ 的范数–弱拓扑下, 集值映射 $\mathcal{D}^+\varphi : V \rightrightarrows X^*$ 在 V 的任意闭子集上是闭图的. 利用该事实和 φ 在 $\bar{x}$ 附近的 Lipschitz 连续性, 由 (6.81) 可导出, 当 x 充分靠近 $\bar{x}$ 时, φ 在经典意义

$$\varphi'(x; z) := \lim_{t \downarrow 0} \frac{\varphi(x + tz) - \varphi(x)}{t}, \quad z \in X.$$

下是方向可导的. 此外, 有表达式

$$\varphi'(x; z) = \min\{\langle x^*, z\rangle \mid x^* \in \mathcal{D}^+\varphi(x)\}, \tag{6.83}$$

其中由于 $\mathcal{D}^+\varphi(x)$ 在 X^* 中的弱闭性, 极小值可达到. 因为 $\mathcal{D}^+\varphi(x)$ 也是凸的, 由 (6.83) 和命题 6.48 的结果有 $\widehat{\partial}^+\varphi(x) \subset \mathcal{D}^+\varphi(x)$. 事实上, 如若不然, 那么将 $x^* \notin \mathcal{D}^+\varphi(x)$ 从凸和依范数拓扑的闭集 $\mathcal{D}^+\varphi(x) \subset X^*$ 中分离开, 就得到与 (6.82) 和 (6.83) 的矛盾. 最后, 由关于基本次梯度极限表示的上次微分形式的定理 2.34,

等式 $\mathcal{D}^+\varphi(x) = \partial^+\varphi(x)$ 和 φ 在 $\bar{x}$ 附近的上正则性由 $\mathcal{D}^+\varphi(\cdot)$ 的闭图性质得到. 这就完成了定理的证明. △

正如上面所提到的, 一致上次可微函数的性质允许在离散逼近最优控制问题中导出具有上次微分横截性条件的 AMP; 参见后面的小节. 与连续时间系统 PMP 中的上次微分横截性条件以及数学规划问题中上次微分最优性条件所需的假设相比, 这对函数和所考虑的空间要求了更多的假设; 参阅 5.1, 5.2 和 6.3 节. 这些对 AMP 所需要的更多限制要求是由于有限差分系统作为当 $N \to \infty$ 时的过程的参数性质所导致的. 在 6.4.3 小节中将看到, 甚至在具有有限维状态空间的自由端点控制问题中的可微费用函数的情形, 导数的连续性对离散逼近序列中 AMP 的正确性也有本质的重要性.

6.4.3 自由端点控制系统的近似最大值原理

本小节致力于轨道右端无端点约束的有限差分系统 (6.77) 序列的最优控制问题的研究. 和连续时间系统情形一样, 离散逼近的自由端点问题和具有约束的问题有着本质的差别. 本小节中主要的肯定性结果是 Banach 空间中具有上次微分横截性条件的自由端点问题的近似最大值原理, 它对一致上次可微费用函数是成立的. 尤其, 这证明了具有连续可微费用函数控制问题的 AMP, 其中伴随系统 (6.78) 的边界/横截性条件可写为经典形式 (6.79). 另一方面, 给出了一个例子表明当费用函数在所考虑的点是可微的但在它的附近不是 $\mathcal{C}^1$ 时, 那么 AMP 不成立. 另外的例子表明 AMP 对非光滑性非常灵敏: 它对具有非光滑动态控制问题, 甚至更为显著的对具有凸非光滑费用函数的良好系统, 是不成立的.

在本小节中, 在满足控制系统 (6.77) 的控制–轨道对 $\{u_N(\cdot), x_N(\cdot)\}$ 上 (当 $N \to \infty$ 时), 考虑离散时间系统最优控制问题序列 (P_N^0):

$$\min \quad J_N[u_N, x_N] := \varphi_0(x_N(b)). \tag{6.84}$$

给定问题 (P_N^0) 的最优解序列 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 作如下常驻假设:

— 控制空间 U 是度量空间, 状态空间 X 是 Banach 空间;

— 存在开集 O 使得 $\bar{x}_N(t) \in O$ $(\forall t \in T_N \cup \{b\})$, 并且当 $N \to \infty$, $x \in O$, $u \in U$ 和 $t \in T_N \cup \{b\}$ 时, f 关于 x Fréchet可微并且 $f(x,u,t)$ 和状态导数 $\nabla_x f(x,u,t)$ 关于 (x,u,t) 连续和一致依范数有界;

— 序列 $\{\bar{x}_N(b)\}$ 属于 X 的一个紧子集.

最后一个假设在有限维空间中是没有多大限制性的: 它可以从确保连续时间控制系统容许轨道的一致有界性的标准条件得到. 在无穷维空间中, 它可由 6.1.1 小节的 (H1) 中的条件导出; 请比照定理 6.13 的证明和其中的参考文献.

现在给出本节主要的肯定性结果.

定理 6.50(具有上次微分横截性条件的自由端点控制问题的 AMP)　设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (P_N^0) 在所作常驻假设下的最优解. 另外假设费用函数 φ_0 在序列 $\{\bar{x}_N(b)\}$ 的极限点附近一致上次可微, 且相应的次梯度集为 $\mathcal{D}^+\varphi_0(x)$. 那么对每个上次梯度序列 $x_N^* \in \mathcal{D}^+\varphi_0(\bar{x}(b))$, 存在 $\varepsilon(t, h_N) \to 0 (N \to \infty)$ 关于 $t \in T_N$ 一致, 使得有近似最大值条件

$$
\begin{aligned}
H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) = & \max_{u \in U} H(\bar{x}_N(t), p_N(t+h_N), u, t) \\
& + \varepsilon(t, h_N), \quad t \in T_N,
\end{aligned} \tag{6.85}
$$

其中每个 $p_N(\cdot)$ 是伴随系统 (6.78) 相应的轨道并且有边界/横截性条件

$$
p_N(b) = -x_N^*, \quad \forall\, N \in \mathbb{N}. \tag{6.86}
$$

更进一步, 如果 X 是自反的, 且 φ_0 在最优点是连续的, 那么 (6.86) 中的结果对任何 $x_N^* \in \widehat{\partial}^+\varphi(\bar{x}_N(b))$ 均成立.

证明　考虑 (P_N^0) 的一列最优解 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 假设轨道 $\bar{x}_N(t)$, 对所有 $N \in \mathbb{N}$ 属于所作假设中固定的一致邻域. 由 φ_0 的一致上次可微性的定义 6.47 有, $\mathcal{D}^+\varphi_0(\bar{x}_N(b)) \neq \varnothing$ 且不等式 (6.81) 对所有 $x_N^* \in \mathcal{D}^+\varphi_0(\bar{x}_N(b))$ (当 $N \to \infty$ 时) 成立. 现在任取序列 $x_N^* \in \mathcal{D}^+\varphi_0(\bar{x}_N(b))$, 有

$$
\varphi_0(x) - \varphi_0(\bar{x}_N(b)) \leqslant \langle x_N^*, x - \bar{x}_N(b) \rangle + o(\|x - \bar{x}_N(b)\|), \tag{6.87}
$$

其中当 $x \to x_N(b)$ 时,

$$
\frac{o(\|x - \bar{x}_N(b)\|)}{\|x - \bar{x}_N(b)\|} \to 0
$$

关于 N 一致. 如 (6.86), 记 $p_N(b) := -x_N^*$, 当 $x_N(b)$ 充分靠近 $\bar{x}_N(b)$ 时, 对 (P_N^0) 中的所有容许过程, 由 (6.87) 可导出

$$
J[u_N, x_N] - J[\bar{u}_N, \bar{x}_N] \leqslant -\langle p_N(b), \Delta x_N(b) \rangle + o(\|\Delta x_N(b)\|),
$$

其中 $\Delta x_N(t) := x_N(t) - \bar{x}_N(t)$. 考虑到恒等式

$$
\begin{aligned}
\langle p_N(b), \Delta x_N(b) \rangle = & \sum_{t \in T_N} \langle p_N(t+h_N) - p_N(t), \Delta x_N(t) \rangle \\
& + \sum_{t \in T_N} \langle p_N(t+h_N), \Delta x_N(t+h_N) - \Delta x_N(t) \rangle
\end{aligned}
$$

并且利用 f 关于 x 的光滑性, 由上面的不等式可得

$$\begin{aligned}0 \leqslant J[u_N, x_N] - J[\bar{u}_N, \bar{x}_N] \leqslant & -\sum_{t\in T_N} \langle p_N(t+h_N) - p_N(t), \Delta x_N(t)\rangle \\ & -h_N \sum_{t\in T_N} \langle p_N(t+h_N), \nabla_x f(\bar{x}_N(t), \bar{u}_N(t), t)\Delta x_N(t)\rangle \\ & -h_N \sum_{t\in T_N} \Delta_u H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) \\ & -h_N \sum_{t\in T_N} \eta_N(t) + o(\|\Delta x_N(b)\|), \end{aligned} \tag{6.88}$$

其中余项 $\eta_N(t)$ 为

$$\begin{aligned}\eta_N(t) = \Big\langle & \nabla_x H(\bar{x}_N(t), p_N(t+h_N), u_N(t), t) \\ & -\nabla_x H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t), \Delta x_N(t)\Big\rangle + o(\|\Delta x_N(t)\|).\end{aligned}$$

并且由关于 $\nabla_x f$ 的假设, $o(\|\Delta x_N(t)\|)$ 关于 N 是一致的, 其中的增量 $\Delta_u H$ 类似 6.3.2 小节中对连续时间系统定义的增量.

现在考虑最优控制 $\bar{u}_N(\cdot)$ 的 (单) 针形变分如下:

$$u_N(t) = \begin{cases} v, & 若\ t = \tau, \\ \bar{u}_N(t), & 若\ t \in T_N \setminus \{\tau\}, \end{cases}$$

其中 $v \in U$ 并且当 $N \in \mathbb{N}$ 时, $\tau = \tau(N) \in T_N$. 所有的这些控制对所考虑的离散逼近问题显然是可行的, 它们并没有端点约束. 对应于针形变分的轨道增量满足关系式

$$\Delta x_N(t) = 0 \quad 对 \quad t = a, \cdots, \tau; \quad \|\Delta x_N(t)\| = O(h_N) \quad 对 \quad t = \tau + h_N, \cdots, b.$$

由此并将针形变分 $u_N(\cdot)$ 替换到增量不等式 (6.88) 中, 则有

$$0 \leqslant J[u_N, x_N] - J[\bar{u}_N, \bar{x}_N] \leqslant -h_N \Delta_u H(\bar{x}_N(\tau), p_N(\tau + h_N), \bar{u}_N(\tau), \tau) + o(h_N).$$

由反证法, 从最后一个不等式可直接导出近似最大值条件 (6.85).

为了完成定理的证明, 剩下来对费用函数 φ_0 利用关于一致上次可微性的定理 6.49. 这确保了如果 X 是自反的并且 φ_0 是连续的, 那么当 $N \to \infty$ 时, x_N^* 可取自于横截性条件 (6.86) 中的整个 Fréchet上次梯度集 $\widehat{\partial}^+\varphi_0(\bar{x}(b))$. 证毕. △

注 6.51(离散逼近与连续时间系统的比较) 注意到定理 6.50 的证明与自由端点连续时间系统中的类似; 请比照 6.3.2 小节定理 6.37 及 6.3.1 小节中上次微分形式 (定理 6.38) 的证明. 在连续时间和离散时间中给出的证明都基于利用费用泛

函的增量公式和最优控制的 (单) 针形变分. 从某种意义上讲, 离散逼近问题的证明是连续时间系统中所给证明的简化形式 (当涉及端点约束时当然就不同了; 参见 6.4.4 小节). 另外, 连续时间系统中的结果和证明与离散逼近中的有两个主要区别.

第一, 在连续时间系统的情形中, 有以小参数 ε 作为针形变分长度的可能性, 在连续时间最优控制中, 这确保了小的轨道增量 $\Delta x(t) = O(\varepsilon)$ 且对确切最大值原理的建立是至关重要的. 在离散逼近系统中, 小轨道增量由递减步长 h_N 的长度给出, 它是问题的一个参数但不是变分的参数. 这导致了具有误差和离散化步长一样小的近似最大值条件. 当然, 当 $h_N \nrightarrow 0$ 时, 这个办法就失效了.

第二, 与连续时间问题相比, 涉及离散逼近问题的参数特性要求对费用函数作更多苛刻的一致性假设.

定理 6.50 及其证明的两个推论处理一些重要的费用函数类, 它们是一致上次可微的且有更细致的 AMP 形式. 注意到, 这些结果与定理 6.50 第二部分不同, 对状态空间 X 不要求自反性; 它们在任何 Banach 空间中都是正确的.

推论 6.52(具有光滑费用函数自由端点控制问题的 AMP) 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (P_N^0) 在所作常驻假设下的最优解. 另外假设费用函数 φ_0 在序列 $\{\bar{x}_N(b)\}$ 的极限点附近是连续可微的. 那么对任意 $N \in \mathbb{N}$, 定理 6.50 中的近似极大原理成立并且具有对应于伴随轨道 $p_N(\cdot)$ 的横截性条件 (6.79). 此外, 如果 $\nabla_x f(\cdot, u, t)$ 和 $\nabla\varphi_0(\cdot)$ 在 $\bar{x}_N(\cdot)$ 附近是局部 Lipschitz 的, 关于 $u \in U$, $t \in T_N$ 一致且 $N \to \infty$, 那么在最大值条件 (6.85) 中可取 $\varepsilon(t, h_N) = O(h_N)$.

证明 如上面提到的, 在任意 Banach 空间 X 中, 如果 φ 在 $\bar{x}$ 附近是 $\mathcal{C}^1$ 的, 那么在该点的邻域内有 $\mathcal{D}^+\varphi(x) = \{\nabla\varphi(x)\}$. 易证对光滑函数 φ_0, (6.87) 作为等式成立; 此外, 如果 $\nabla\varphi_0$ 是局部 Lipschitz 的, 那么有 $|o(\eta)| \leqslant \ell\eta^2$. 更进一步, 在定理的证明中对余项 $\eta_N(\cdot)$ 中的“o”项, $\nabla_x f(\cdot, u, t)$ 的 Lipschitz 假设蕴涵着

$$o(\|\Delta x_N(t)\|) = O(\|\Delta x_N(t)\|^2)$$

关于 N 一致. 这在近似最大值条件 (6.85) 中就得到 $\varepsilon(t, h_N) = O(h_N)$. △

推论 6.53(具有凹费用函数自由端点控制问题的 AMP) 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (P_N^0) 在所作常驻假设下的最优解, 另外假设费用函数 φ_0 在包含所有 $\bar{x}_N(b)$ 的某开集上是凹的, 那么沿着次梯度 $x_N^* \in \partial^+\varphi_0(\bar{x}_N(b))$ 的每个序列, 定理 6.50 中的近似极大原理成立. 此外, 如果 $\nabla_x f(\cdot, u, t)$ 在 $\bar{x}_N(\cdot)$ 附近是局部 Lipschitz 的, 关于 $u \in U$, $t \in T_N$ 一致且 $N \to \infty$, 那么在最大值条件 (6.85) 中可取 $\varepsilon(t, h_N) = O(h_N)$.

证明 在任意 Banach 空间中, 对凹连续函数有 $\mathcal{D}^+\varphi(x) = \partial^+\varphi(x)$. 更进一步, 在凹性的假设下, 在不等式 (6.87) 中有 $o(\|x - \bar{x}_N(b)\|) \equiv 0$. 将它与推论 6.52 证明中 $\eta_N(\cdot)$ 的估计结合起来, 在 (6.85) 中就可得到 $\varepsilon(t, h_N) = O(h_N)$. △

现在考虑反例, 即如果定理 6.50 中的一些假设不满足, 那么 AMP 不再成立. 下面所有的例子在具有非凸速率集的有限维控制系统中给出. 例 6.54 表明 AMP 对预期的下次微分形式 (像连续时间控制系统的最大值原理那样) 不成立, 甚至在线性动态上极小化凸函数的最简单非光滑情形亦如此.

例 6.54(具有非光滑和凸极小函数线性控制系统的 AMP 可能不成立) 在无端点约束的线性系统上, 存在极小化非光滑和凸费用函数的一维控制问题使得 AMP 不成立.

证明 考虑下面一维离散时间系统最优控制问题序列 (P_N^0), $N \in \mathbb{N}$:

$$\begin{cases} \min & \varphi(x_N(1)) := |x_N(1) - \vartheta| \\ \text{s.t.} & x_N(t+h_N) = x_N(t) + h_N u_N(t), \quad x_N(0) = 0, \\ & u_N(t) \in U := \{0,1\}, \quad t \in T_N := \{0, h_N, \cdots, 1-h_N\}, \end{cases} \tag{6.89}$$

其中 ϑ 是小于 1 的正无理数, 它的选取将在下面具体给出. (6.89) 中的动态是最简单 ODE 控制系统 $\dot{x} = u$ 的离散化. 注意到, 因为 ϑ 是无理数并且 h_N 是有理数, 那么对于 (6.89) 的最优轨道的端点有 $\bar{x}_N(1) \neq \vartheta$ $(N \in \mathbb{N})$. 然而对连续时间系统的最优解显然有 $\bar{x}(1) = \vartheta$. 明显也有, 对充分小的步长 h_N, (6.89) 的最优控制既不是 $u_N(t) \equiv 0$, 也不是 $u_N(t) \equiv 1$, 但它至少有一个控制转换点.

假设对某子列 $N_k \to \infty$ 有 $\bar{x}_{N_k}(1) > \vartheta$; 不失一般性, 令 $\{N_k\} = \mathbb{N}$. 要证明, 在此情形, 近似最大值条件在满足 $\bar{u}_N(t) = 1$ 的点 $t \in T_N$ 不成立. 事实上, 对 Hamilton-Pontryagin 函数和该问题的伴随轨道有

$$H(\bar{x}_N(t), p_N(t+h_N), u) = p_N(t+h_N)u \quad \text{和} \quad p_N(t) \equiv -1,$$

这是因为沿着 (6.89) 的最优解有 $\bar{x}_N(1) > \vartheta$. 因而在控制转换点 $s \in T_N$ 有

$$\max_{u \in U} H(\bar{x}_N(t), p_N(t+h_N), u) = 0, \quad t \in T_N,$$

$$\text{但} \quad H(\bar{x}_N(s), p_N(s+h_N), \bar{u}_N(s)) = -1,$$

其中 $\bar{u}(s) = 1$ 与 h_N 无关.

现在具体选取 (6.89) 中的 ϑ 以确保沿着自然数的某子列有 $\bar{x}_N(1) > \vartheta$. 取无理数 $\vartheta \in (0,1)$ 使其十进制小数表示包含集合 $\{5,6,7,8,9\}$ 中的无穷多数字; 例如, $\vartheta = 0.676676667\cdots$, 那么可断言 $\bar{x}_N(1) > \vartheta$. 事实上, 令 $h_N := 10^{-N}$, 这是 (6.89) 中要求的 $h_N = N^{-1}$ 的一个子列. 易见, 在该情形中, 在 $t = 1$ 的所有可达点集包含介于 0 和 1 之间的所有有理数, 其中小数部分的十进制表示恰好有 N 位数. 特别地, 对 $N = 3$, 该集合为 $\{0, 0.001, 0.002, \cdots, 0.999, 1\}$. 因此, 由 ϑ 的构造, 可达集最靠近 ϑ 的点大于 ϑ 并且这个点一定是最优轨道的端点 $\bar{x}_N(1)$. △

下面一个例子, 作为例 6.54 的补充, 表明了即使对可微但不连续可微费用函数的问题, AMP 也可能不成立.

例 6.55(具有可微但不是 $\mathcal{C}^1$ 费用函数线性控制系统的 AMP 可能不成立) 在无端点约束的线性系统上, 存在极小化 Fréchet可微但不连续可微费用函数的一维控制问题使得 AMP 不成立.

证明 考虑 (6.89) 中同样的控制系统且构造极小化函数 $\varphi(x)$ 满足上面的要求. 设 $\psi(x)$ 为 $\mathcal{C}^1$ 函数并且有性质:

$$\psi(x) \geqslant 0, \quad \psi(x) = \psi(-x), \quad \psi(x) \equiv 0, \quad \text{若} \quad |x| > 2,$$
$$|\nabla\psi(x)| \leqslant 1, \quad \text{对所有 } x, \quad \nabla\psi(-1) = \vartheta > 0.$$

定义费用函数 $\varphi(x)$ 为

$$\varphi(x) := \left(x - \frac{1}{9}\right)^2 + \sum_{n=1}^{\infty} 10^{-2n-3}\psi\left(10^{2n+3}\left(x - \sum_{k=1}^{n} 10^{-k}\right) - 1\right),$$

它在除了 $x = \dfrac{1}{9}$ 外的每个点附近是连续可微的, 在 $x = \dfrac{1}{9}$ 处仅是可微的并且达到绝对最小值. 像例 6.54, 令 $h_N := 10^{-N}$, 那么点 $x = \dfrac{1}{9}$ 由离散化不可能达到. 不难验证对每个 N, 最优轨道 $\bar{x}_N(\cdot)$ 的端点

$$\bar{x}_N(1) = \sum_{k=1}^{N} 10^{-k} \quad \text{并且} \quad \nabla\varphi(\bar{x}_N(1)) = \vartheta + \varepsilon_N,$$

其中当 $N \to \infty$ 时, $\varepsilon_N \downarrow 0$. 像例 6.54 中那样, 利用同样的最优控制序列, 于是有

$$H(\bar{x}_N(t), p_N(t + h_N), u) \equiv -\vartheta u + O(\varepsilon_N),$$

并且近似最大值条件 (6.85) 在控制转换端点不成立, 其中 $\bar{u}_N(t) = 1$. △

本小节最后一个例子考虑非光滑动态系统. 实际上考虑类似于一微控制系统上的极小化积分泛函的有限差分, 它等价于二维 Mayer 最优控制问题. 在这个问题中, 离散的"被积函数"关于状态变量 x 是非光滑的; 沿着最优过程 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 它关于 x 是连续可微的但关于 N 不是一致的. 因而下面的例子表明了在包含所有最优轨道 $\bar{x}_N(\cdot)$ 的开集上, 关于 f 的一致光滑性假设对 AMP 是本质的.

例 6.56(违反 AMP 的非光滑动态控制问题) 在无端点约束的一维线性控制系统上, 被积函数关于控制变量是线性的且关于状态变量是凸和非光滑的, 那么离散逼近极小化积分泛函问题的 AMP 不成立. 此外, 在该问题中对所有 $N \in \mathbb{N}$, 沿着离散逼近 (P_N^0) 的最优解序列, 被积函数关于状态变量是 $\mathcal{C}^1$ 的但关于 N 不是一致的.

证明 首先考虑下面的 Bolza 型连续时间最优控制问题:

$$\begin{cases} \min & J[u,x] := \displaystyle\int_0^b (u(t) + |x(t) - t^2/2|)\mathrm{d}t \\ \text{s.t.} & \dot{x} = tu, \quad x(0) = 0, \\ & u(t) \in U := \{1, c\}, \quad 0 \leqslant t \leqslant b, \end{cases}$$

其中终端时间 b 和数 $c > 1$ 将在下面给出. 显然该问题的最优控制是 $\bar{u}(t) \equiv 1$ 并且相应的最优轨道为 $\bar{x}(t) = t^2/2$.

由离散化, 可得到有限差分控制问题序列:

$$\begin{cases} \min & J[u_N, x_N] := h_N \displaystyle\sum_{t \in T_N} (u_N(t) + |x_N(t) - t^2/2|) \\ \text{s.t.} & x_N(t + h_N) = x_N(t) + h_N t u_N(t), \quad x_N(0) = 0, \\ & u_N(t) \in U := \{1, c\}, \quad t \in T_N := \{0, \cdots, (N-1)h_N\}. \end{cases} \tag{6.90}$$

首先证明如果步长 h_N 充分小并且数 (b, c) 适当选取, 那么 $\bar{u}_N(t) \equiv 1$ 仍然是 (6.90) (唯一) 的最优控制. 容易验证相应的轨道 $\bar{x}(\cdot)$ 为

$$\bar{x}_N(t) = \frac{t^2}{2} - \frac{th_N}{2}, \quad \forall N \in \mathbb{N}.$$

那么费用泛函在 $\bar{u}_N(\cdot)$ 的值 $\bar{J}_N$ 等于

$$\bar{J}_N = b + h_N^2 \sum_{t \in T_N} \frac{t}{2} = b + \frac{b^2 h_N}{4} + o(h_N).$$

如果在某点 $t \in T_N$ 用 $u_N(t) = c$ 替换 $u_N(t) = 1$, 那么和式 $h_N \sum\limits_{t \in T_N} u_N(t)$ 的增量等于 $(c-1)h_N$. 于是对 (6.90) 的与 $\bar{u}_N(t) \equiv 1$ 不同的任意可行控制 $u_N(t)$, 费用泛函相应的值为

$$\begin{aligned} J[u_N, x_N] &= h_N \sum_{t \in T_N} u_N(t) + h_N \sum_{t \in T_N} |x_N(t) - t^2/2| \\ &> h_N \sum_{t \in T_N} u_N(t) \geqslant b + (c-1)h_N. \end{aligned}$$

与 $\bar{J}_N$ 相比, 可得到, 如果 $b^2/4 < c - 1$ 且 N 充分大, 那么控制 $\bar{u}_N(t) \equiv 1$ 的确是 (6.90) 的最优控制.

最后证明对 $b > 2$ 和 $c > b^2/4 + 1$ (例如, 对 $b = 3$ 和 $c = 4$), 最优控制序列 $\bar{u}_N(t) \equiv 1$ 在所有充分靠近 $t = b/2$ 的点 $t \in T_N$ 不满足近似最大值条件 (6.85). 在对应于最优控制并且具有满足 (6.78) 的伴随轨道 $p_N(t)$ 的最优轨道 $\bar{x}_N(t)$ 上, 计算

关于 $t \in T_N$ 和 $u \in U$ 的 Hamilton-Pontryagin 函数. 将 (6.90) 约化为标准 Mayer 型来考虑, 由上面关于 $\bar{x}_N(t)$ 的公式, 对所有 $t \in T_N$, $\bar{x}_N(t) < t^2/2$, 有

$$\begin{aligned} H(\bar{x}_N(t), p_N(t+h_N), u, t) &= tp_N(t+h_N)u - u - |\bar{x}_N(t) - t^2/2| \\ &= (tp_N(t+h_N) - 1)u + (\bar{x}_N(t) - t^2/2), \end{aligned}$$

其中 $p_N(t)$ 满足方程

$$p_N(t) = p_N(t+h_N) + h_N, \quad p_N(b) = 0,$$

它的解为 $p_N(t) = b - t$. 因此有

$$\begin{aligned} H(\bar{x}_N(t), p_N(t+h_N), u, t) &= (t(b-t+h_N) - 1)u + O(h_N) \\ &= (-t^2 + bt - 1)u + O(h_N). \end{aligned}$$

如果判别式 $b^2 - 4$ 是正的, 那么乘子 $-t^2 + bt - 1$ 在 $t = b/2$ 的邻域内是正的. 因而, 如果 h_N 充分小, 那么 $u = c$, 而不是 $u = 1$, 在 $t = b/2$ 处为 Hamilton-Pontryagin 函数提供了最大值. 证毕. △

在本小节最后, 给出定理 6.50 在时间区间长度 $b - a$ 和步长 h_N 不可公度的一般情形的一个修改; 与此不同, 在定理 6.50 中, 对任意 $N \in \mathbb{N}$ 有 $b - a = Nh_N$. 该修改对 AMP 扩展到 6.4.5 小节中时滞系统的有限差分逼近尤其重要. 为了简单起见, 利用记号

$$f(x_N, u_N, t) := f(x_N(t), u_N(t), t).$$

给定时间区间 $[a, b]$, 在 $[a, b]$ 上定义格点 T_N 为

$$\begin{aligned} &T_N := \{a, a+h_N, \cdots, b - \widetilde{h}_N - h_N\}, \\ &\text{其中} \quad h_N := \frac{b-a}{N} \quad \text{和} \quad \widetilde{h}_N := b - a - h_N \left[\frac{b-a}{h_N}\right], \end{aligned}$$

这里 $[z]$ 表示小于或等于实数 z 的最大整数. 于是修改后的离散逼近问题 $(\widetilde{P}_N^0)$ 可写为

$$\begin{cases} \min & J[u_N, x_N] := \varphi_0(x_N(b)) \\ \text{s.t.} & x_N(t+h_N) = x_N(t) + h_N f(x_N, u_N, t), \quad t \in T_N, \quad x_N(a) = x_0 \in X, \\ & x_N(b) = x_N(b - \widetilde{h}_N) + \widetilde{h}_N f(x_N, u_N, b - \widetilde{h}_N), \\ & u_N(t) \in U, \quad t \in T_N. \end{cases}$$

定理 6.57(不可公度问题的 AMP) 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 $(\widetilde{P}_N^0)$ 在所作常驻假设下的最优解. 另外假设 φ_0 在序列 $\{\bar{x}_N(b)\}$ $(N \in \mathbb{N})$ 的极限点附近是一

致上次可微的. 那么对每个上次梯度序列 $x_N^* \in \mathcal{D}^+\varphi_0(\bar{x}_N(b))$, 存在 $\varepsilon(t,h_N) \to 0$ $(N \to \infty)$, 关于 t 一致, 使得近似最大值条件

$$H(\bar{x}_N, p_N, \bar{u}_N, t) = \max_{u \in U} H(\bar{x}_N, p_N, u, t) + \varepsilon(t, h_N)$$

对所有 $t \in \widetilde{T}_N := T_N \cup \{b - \widetilde{h}_N\}$ 成立, 其中 Hamilton-Pontryagin 函数定义为

$$H(\bar{x}_N, p_N, u, t) := \begin{cases} \langle p_N(t+h_N), f(\bar{x}_N, u, t)\rangle, & \text{如果 } t \in T_N, \\ \langle p_N(t), f(\bar{x}_N, u, t - \widetilde{h}_N)\rangle, & \text{如果 } t = b - \widetilde{h}_N, \end{cases}$$

并且每个 $p_N(\cdot)$ 满足伴随系统

$$\begin{cases} p_N(t) = p_N(t+h_N) + h_N \nabla_x f(\bar{x}_N, \bar{u}_N, t)^* p_N(t+h_N), & t \in T_N, \\ p_N(b - \widetilde{h}_N) = p_N(b) + \widetilde{h}_N \nabla_x f(b - \widetilde{h}_N, \bar{x}_N, \bar{u}_N, t)^* p_N(b) \end{cases}$$

以及有横截性条件 $p_N(b) = -x_N^*$. 更进一步, 类似定理 6.50 的第二部分及推论 6.52 和推论 6.53 的一些具体化也得到满足.

证明 类似定理 6.50 及其推论的证明, 其中利用如下修改的极小化泛函增量公式

$$\begin{aligned} 0 \leqslant J[u_N, x_N] - J[\bar{u}_N, \bar{x}_N] &\leqslant -\langle p_N(b), \Delta x_N(b)\rangle + o(\|\Delta x_N(b)\|) \\ &= -\sum_{t \in T_N} \langle p_N(t+h_N) - p_N(t), \Delta x_N(t)\rangle \\ &\quad - \langle p_N(b) - p_N(b - \widetilde{h}_N), \Delta x_N(b - \widetilde{h}_N)\rangle \\ &\quad - h_N \sum_{t \in T_N} \langle p_N(t+h_N), \nabla_x f(\bar{x}_N, \bar{u}_N, t)\Delta x_N(t)\rangle \\ &\quad - \widetilde{h}_N \langle p_N(b), \nabla_x f(\bar{x}_N, \bar{u}_N, b - \widetilde{h}_N)\Delta x_N(b - \widetilde{h}_N)\rangle \\ &\quad - h_N \sum_{t \in \widetilde{T}_N} \Delta_u H(\bar{x}_N, p_N, \bar{u}_N) + h_N \sum_{t \in \widetilde{T}_N} \eta_N(t) + o(\|\Delta x_N(b)\|), \end{aligned}$$

其中 $\Delta_u H$ 和 $\eta_N(t)$ 类似于无时滞问题中的定义. 将伴随轨道代入该公式中并利用最优控制的针形变分, 则得到定理的结论. △

6.4.4 端点约束下的近似最大值原理: 肯定和否定的陈述

本小节考虑具有端点约束最优控制问题的离散逼近. 这里主要的目标是在适当的假设下阐述离散逼近问题的近似最大值原理, 并且阐明这些假设对它的正确性是否是本质的; AMP 的证明在 6.4.5 小节中给出.

在构造离散逼近时, 自然要扰动端点约束并且考虑下面的离散时间系统最优控制问题序列 (P_N):

$$\begin{cases} \min \quad J[u_N, x_N] := \varphi_0(x_N(b)) \\ \text{s.t.} \quad x_N(t+h_N) = x_N(t) + h_N f(x_N(t), u_N(t), t), \quad x_N(a) = x_0 \in X, \\ \qquad u_N(t) \in U, \quad t \in T_N := \{a, a+h_N, \cdots, b-h_N\}, \\ \qquad \varphi_i(x_N(b)) \leqslant \gamma_{iN}, \quad i = 1, \cdots, m, \\ \qquad |\varphi_i(x_N(b))| \leqslant \xi_{iN}, \quad i = m+1, \cdots, m+r, \\ \qquad h_N := \dfrac{b-a}{N}, \quad N = 1, 2, \cdots, \end{cases}$$

其中对所有 i, 当 $N \to \infty$ 时, $\gamma_{iN} \to 0$ 且 $\xi_{iN} \downarrow 0$. 本小节的主要结果表明, 在对初始数据所作的标准假设下, 对于具有不等式约束的任意扰动 (特别地, 可令 $\gamma_{iN} = 0$) 和具有与离散步长相吻合的等式约束相容扰动问题 (P_N) 的最优控制恰当序列, AMP 成立. 然后说明所提到的恰当性和相容性要求对 AMP 的正确性是本质的, 同时也导出具有非光滑费用和不等式约束函数问题 AMP 的一个适当上次微分形式.

在整个的这一小节中, 保持 6.4.3 小节在初始数据上所作的常驻假设, 另外假设状态空间 X 是有限维的, 这在下面的证明中用到. 与常规的矩阵乘法记号一道, 利用协定

$$\prod_{j}^{k=i} A_k := \begin{cases} A_i A_{i-1} \cdots A_j, & \text{如果 } i \geqslant j, \\ I, & \text{如果 } i = j-1, \\ 0, & \text{如果 } i < j-1, \end{cases}$$

其中 i, j 为任意整数并且依惯例, I 表示单位矩阵.

像连续时间系统的情形一样, 具有端点约束问题 (P_N) 的 AMP 的证明与具有自由端点问题的相比, 有着本质的差别且更复杂. 回顾 6.3.3 小节中具有不等式端点约束的连续时间系统的定理 6.37 的证明, 注意到该证明的关键部分是引理 6.44, 它证明了 (6.74) 中的线性化映像集 S 是凸的并且与禁止点的集合不相交. 这些事实当然是由于时间的连续性并反映了连续时间控制系统的隐含凸性. 注意, 所提到的 (6.74) 中的映像集 S 是由最优控制的多针形变分生成的, 它在 (6.82) 中的构造本质上基于时间的连续性.

下面建立类似 (P_N) 中控制系统隐含凸性性质的某种有限差分, 它涉及一些由最优控制单针形变分生成的线性化映像集 S_N 的凸包. 下面证明对这些凸包作小的平移 (不超过 $o(h_N)$) 后, 当 $N \to \infty$ 时, 与禁止点的集合不相交. 由凸集分离定理, 这基本上导致了在不等式型端点约束下问题 (P_N) 的近似最大值原理并且具有适当的扰动互补松弛条件.

这种方法 (以及 6.3.4 小节中构造的任何有限差分版本) 并不适用于具有任意扰动的等式约束 (特别地, 当 $\xi_N = 0$ 时) 问题 (P_N), 此时 AMP 一般并不成立. 然

而, 上面提到的互补松弛条件允许导出具有适当扰动的等式约束 (约化为不等式约束) 问题 (P_N) 的 AMP 的一种自然形式.

在阐述本小节主要结果之前, 首先引入一个针对有限差分控制问题序列的重要概念.

定义 6.58(离散逼近中控制的恰当性) 设 $d(\cdot,\cdot)$ 表示问题 (P_N) 中控制空间 U 中的距离. 称 (P_N) 中离散时间控制序列 $\{u_N(\cdot)\}$ 是恰当的, 如果对每个单增的自然数子列 $\{N\}$ 和每个网点序列 $\tau_{\theta(N)} \in T_N$ 满足

$$\tau_{\theta(N)} = a + \theta(N)h_N, \text{ 当 } \theta(N) = 0, \cdots, N-1 \text{ 时, 并且 } \tau_{\theta(N)} \to t \in [a,b],$$

那么当 $N \to \infty$ 时, 并且对任何自然常数 q, 下面性质中的一个成立:

$$d(u_N(\tau_{\theta(N)}), u_N(\tau_{\theta(N)+q})) \to 0, \text{ 或者 } d(u_N(\tau_{\theta(N)}), u_N(\tau_{\theta(N)-q})) \to 0.$$

在离散逼近问题中, 可行控制序列的恰当性概念是连续时间系统分段连续性的有限差分版本. 结果, 在具有非凸速率约束系统的离散逼近中, 最优控制序列经常不是恰当的, 这导致了不等式约束标准问题的 AMP 不成立. 注意到, 在自由端点问题中, 恰当性假设对于 AMP 的正确性不是必需的; 参见定理 6.50.

现在可以阐述由光滑函数描述的端点约束控制问题 (P_N) 的 AMP 了.

定理 6.59(具有光滑端点约束控制问题的 AMP) 在所作的常驻假设下, 对所有 $N \in \mathbb{N}$, 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为 (P_N) 的最优解. 另外假设所有的函数 φ_i, $i = 0, \cdots, m+r$, 在 $\{\bar{x}_N(b)\}$ 的极限点附近是连续可微的, 并且

(a) 最优控制序列 $\{\bar{u}_N(\cdot)\}$ 是恰当的;

(b) 相容性条件 (6.80) 对所有等式约束扰动 ξ_{iN} 成立.

那么存在数 $\{\lambda_{iN} | i = 0, \cdots, m+r\}$ 满足

$$\lambda_{iN}(\varphi_i(\bar{x}_N(b)) - \gamma_{iN}) = O(h_N), \quad i = 1, \cdots, m, \tag{6.91}$$

$$\lambda_{iN} \geqslant 0, \ i = 0, \cdots, m, \quad \sum_{i=0}^{m+r} \lambda_{iN}^2 = 1, \tag{6.92}$$

使得近似最大值条件 (6.85) 得到满足, 并且当 $N \to \infty$ 时, $\varepsilon_N(t, h_N) \to 0$ 对 $t \in T_N$ 一致, 其中 $p_N(t)$, $t \in T_N \cup \{b\}$ 是伴随系统 (6.78) 的相应轨道并且具有端点横截性条件

$$p_N(b) = -\sum_{i=0}^{m+r} \lambda_{iN} \nabla \varphi_i(\bar{x}_N(b)). \tag{6.93}$$

这里将这个主要定理的证明放到 6.4.5 小节. 现在给出两个反例来表明, 在其他假设成立的条件下, 恰当性和相容性条件对 AMP 的正确性是本质的. 例 6.60 涉及定义 6.58 中的恰当性条件.

例 6.60(在光滑控制问题中没有恰当条件则 AMP 可能不成立) 存在具有一个不等式约束的二维线性控制问题, 使得离散逼近序列中的最优控制不是恰当的并且不满足近似最大值原理.

证明 考虑具有如下形式的一个线性连续时间最优控制问题 (P), 其中二维状态 $x=(x_1,x_2)\in\mathbb{R}^2$:

$$\begin{cases} \min & \varphi(x(1)) := -x_1(1) \\ \text{s.t.} & \dot{x}_1 = u, \quad \dot{x}_2 = x_1 - ct, \quad x_1(0)=x_2(0)=0, \\ & u(t)\in U := \{0,1\}, \quad 0\leqslant t\leqslant 1, \\ & x_2(1)\leqslant -\dfrac{c-1}{2}, \end{cases}$$

其中 $c>1$ 是一个给定的常数. 注意到, 该问题唯一"讨厌"的特性是控制集 $U=\{0,1\}$ 是非凸的. 于是可行速率集 $f(x,U,t)$ 也是非凸的. 很明显, $\bar{u}(t)\equiv 1$ 是问题 (P) 的唯一最优解, 并且相应的最优轨道为 $\bar{x}_1(t)=t$, $\bar{x}_2=-\dfrac{c-1}{2}t^2$. 此外, 不等式约束是积极/起作用的, 因为 $\bar{x}_2(1)=-\dfrac{c-1}{2}$.

现在离散化该问题并且步长 $h_N := \dfrac{1}{2N}$, $N\in\mathbb{N}$. 为了记号的方便, 下面略去指标 N. 于是相应于上面问题 (P) 的离散逼近问题 (P_N) 可写为

$$\begin{cases} \min & \varphi(x(1)) = -x_1(1) \\ \text{s.t.} & x_1(t+h) = x_1(t)+hu(t), \quad x_1(0)=0, \\ & x_2(t+h) = x_2(t)+h(x_1(t)-ct), \quad x_2(0)=0, \\ & u(t)\in\{0,1\}, \quad t\in\{0,h,\cdots,1-h\}, \\ & x_2(1)\leqslant -\dfrac{c-1}{2}+h^2, \end{cases}$$

即在 (P_N) 的约束扰动中令 $\gamma_N := h_N^2$.

为此, 计算 (P_N) 中相应于 $u(t)\equiv 1$ 的轨道. 易见, 对该 $u(\cdot)$, 有 $x_1(t)=t$. 为了计算 $x_2(t)$, 注意到

$$[x(t+h)=x(t)+ht,\ x(0)=0] \Longrightarrow x(t) = \frac{t^2}{2}-\frac{th}{2}.$$

事实上, 直接计算可得

$$x(t) = h\sum_{\tau=0}^{t-h} = [\text{令 } \tau = kh] = h^2\sum_{k=0}^{\frac{t}{h}-1} k = h^2\frac{\frac{t}{h}(\frac{t}{h}-1)}{2} = \frac{t^2}{2}-\frac{th}{2}.$$

因此, 对 (P_N) 中相应于 $u(t)\equiv 1$ 的 $x_2(t)$, 有

$$x_2(t) = h\sum_{\tau=0}^{t-h}(\tau - c\tau) = -\frac{c-1}{2}t^2+\frac{c-1}{2}ht.$$

由这个计算可看到, 对充分小的 h, $x_2(1)$ 不再满足端点约束, 于是对所有靠近零的 h, $u(t)\equiv 1$ 不是问题 (P_N) 的可行控制. 这意味着对小的 h, (P_N) 的最优控制显然存在, 并且至少有一个转换点 s 使得 $\bar{u}(s)=0$. 因此对应于端点 $x_1(1)$ 的极大值将小于或等于 $1-h$. 令

$$u(t):=\begin{cases}1, & \text{如果 } t\neq s,\\ 0, & \text{如果 } t=s\end{cases}$$

并且对 (P_N) 中依赖于 h 和 s 的相应轨道, 验证公式

$$x_2(t)=\begin{cases}-\dfrac{c-1}{2}t^2+\dfrac{c-1}{2}ht, & t\leqslant s,\\ -\dfrac{c-1}{2}t^2+\dfrac{c-1}{2}ht-h(t-s)+h^2, & t\geqslant s+h.\end{cases}$$

只需验证该公式的第二部分. 对 $t\geqslant s+h$, 为了计算 $x_2(t)$, 将 $x_1(t)=t-h$ 代入 (P_N) 中的离散系统. 易见, 与 $u(t)\equiv 1$ 时的情形相比较, 增量 $\Delta x_2(t)$ 为

$$h\sum_{\tau=s+h}^{t-h}(-h)=-h(t-h-s)=-h(t-s)+h^2,$$

这就证明了上面关于 $x_2(t)$ 的公式.

现在将上述控制具体化, 对所有 N, 令 $c=2$, $s=0.5$, 即考虑离散时间函数

$$\bar{u}(t):=\begin{cases}1, & \text{如果 } t\neq 0.5,\\ 0, & \text{如果 } t=0.5.\end{cases}$$

注意到, 由于 $h_N:=\dfrac{1}{2N}$, 对所有 N, 点 $t=0.5$ 属于格点 T_N. 进一步注意到这些控制序列不满足定义 6.58 中的恰当性性质. 由上面关于 $x_2(t)$ 的公式可得, 只要 $N\in\mathbb{N}$, 相应的轨道遵循 (P_N) 中的端点约束, 因为 $\bar{x}_2(1)=-\dfrac{1}{2}+h^2$. 此外, 由所给的计算易得, 对任何 N, 控制 $\bar{u}(t)$ 是问题 (P_N) 的最优控制.

接下来证明该最优控制序列 $\bar{u}(\cdot)$ 在转换点不满足近似最大值条件 (6.85). 事实上, 问题 (P_N) 的伴随系统 (6.78) 为

$$p(t)=p(t+h)+h\nabla_x f(\bar{x}_1,\bar{x}_2,\bar{u},t)^*p(t+h),$$

其中 Jacobi 矩阵 $\Delta_x f$ 及其伴随/转置为

$$\nabla_x f=\begin{pmatrix}0 & 0\\ 1 & 0\end{pmatrix},\quad \nabla_x f^*=\begin{pmatrix}0 & 1\\ 0 & 0\end{pmatrix}.$$

因此有伴随轨道

$$p_1(t) = p_1(t+h) + hp_2(t+h), \quad \text{并且 } p_2(t) \equiv \text{常数},$$

其中 (p_1, p_2) 满足横截性条件 (6.93). 且符号和非平凡条件 (6.92) 可写为

$$p_1(1) = \lambda_0,\ p_2(1) = -\lambda_1; \quad \lambda_0 \geqslant 0,\ \lambda_1 \geqslant 0,\ \lambda_0^2 + \lambda_1^2 = 1.$$

这蕴涵着 $p_1(t)$ 是一个线性非减函数, 相应的 Hamilton-Pontryagin 函数等于

$$H(x(t), p(t+h), u(t)) = p_1(t+h)u(t) + \text{不依赖 } u \text{ 的项}.$$

由该表达式并考虑到对所有除了 $t = 0.5$ 外的 t, 最优控制 $\bar{u}(t) = 1$, 那么可断言近似最大值条件 (6.85) 成立仅当 $p_1(t)$ 是非负或者除 $t = 0.5$ 外处处趋于零. 注意到由 $p_1(t) \equiv 0$ 可得 $\lambda_1 = \lambda_2 = 0$, 这与非平凡条件相矛盾. 于是 $p_1(t)$ 除了在 $t = 0$ 外是正数. 因此, 当 $h \downarrow 0$ 时, 有一个转换点不趋于零的控制序列, 在该点不可能满足近似最大值条件. 这就证明了上面构造的问题 (P_N) 的最优控制序列的 AMP 不成立. △

基于上面的思想, 可以构造许多这类例子, 主要步骤如下. 取一个具有起作用的不等式约束和非凸容许速率集 $f(x, U, t)$ 的连续时间问题. 通常经过离散化以后, “原来的” 最优控制在离散逼近中不再是容许的, 并且在离散时间问题序列中 “新的” 最优控制有一个奇异转换点 (因而使得最优控制序列不再恰当), 其中的近似最大值条件不满足.

下面一个例子表明如果等式约束没有离散化步长相容扰动, 那么对于具有等式端点约束的连续时间系统离散逼近最优控制的恰当序列 AMP 可能不成立.

例 6.61(没有相容扰动的等式约束 AMP 可能不成立) 存在具有一个等式型线性端点约束的二维线性控制问题, 使得如果离散逼近序列中的恰当最优控制没有相容扰动, 那么它不满足 AMP.

证明 首先考虑如下具有等式型端点约束的二维系统最优控制问题:

$$\begin{cases} \min & \varphi_0(x(1)) := x_2(1) \\ \text{s.t.} & \dot{x} = u, \quad t \in T := [0,1], \quad x(0) = 0, \\ & u(t) \in U := \left\{(0,0), (0,-1), (1,-\sqrt{2}), (-\sqrt{2},-3)\right\}, \\ & \varphi_1(x(1)) := x_1(1) = 0, \end{cases}$$

其中 $x = (x_1, x_2) \in \mathbb{R}^2$ 并且 $u = (u_1, u_2) \in \mathbb{R}^2$. 易见, 该线性问题已是尽可能的标准和简单, 唯一例外的是控制区域 U 的非凸性.

以定理 6.59 中的标准方式, 取零扰动的端点约束, 即 $\xi_N = 0$, 来构造离散逼近问题序列 (P_N), 即

$$\begin{cases} \min & \varphi_0(x_N(1)) = x_{2N}(1) \\ \text{s.t.} & x_N(t+h_N) = x_N(t) + h_N u_N(t), \quad x_N(0) = 0 \in \mathbb{R}^2, \\ & u(t) \in U, \quad t \in T_N := \{0, h_N, \cdots, 1-h_N\}, \\ & \varphi_1(x_N(1)) = x_{1N}(1) = 0 \quad \text{并且} \quad h_N = N^{-1}, \quad N \in \mathbb{N}. \end{cases}$$

容易验证, 问题 (P_N) 的唯一最优解为

$$\bar{u}_N(t) = (0, -1), \quad \bar{x}_N(t) = (0, -t), \quad \forall t \in T_N \cup \{1\}, \quad N \in \mathbb{N}.$$

这给出了费用泛函的最小值 $\bar{J}_N = -1$. 注意到, 序列 $\{\bar{u}_N(\cdot)\}$ 在定义 6.58 的意义下显然是恰当的. 满足横截性条件 (6.93) 的伴随系统 (6.78) 的相应轨道 $p_N(\cdot)$ 为

$$p_N(t) = (-\lambda_{1N}, -\lambda_{0N}) \quad (\forall t \in T_N \cup \{1\}),$$

其中乘子 $(\lambda_{0N}, \lambda_{1N})$ 的符号和非平凡条件 (6.92) 可写为

$$\lambda_{0N} \geqslant 0, \quad \lambda_{0N}^2 + \lambda_{1N}^2 = 1 \quad (\forall N \in \mathbb{N}).$$

更进一步, 对每个 $N \in \mathbb{N}$, 沿着 $\bar{x}_N(\cdot)$ 和相应的伴随轨道 $p_N(\cdot)$, 离散时间系统的 Hamilton-Pontryagin 函数可约化为

$$H_N(u, t) = -\lambda_{1N} u_1 - \lambda_{0N} u_2, \quad t \in T_N,$$

这对最优控制有 $H_N(\bar{u}_N) = \lambda_{0N}$.

现在来证明估计

$$\delta_N := \max\{H_N(u) | u \in U\} - H_N(\bar{u}_N) \geqslant 1 \quad (\forall N \in \mathbb{N}).$$

这表明上面的问题序列 (P_N) 违反了近似最大值条件 (6.85). 为此, 考虑乘子 $(\lambda_{0N}, \lambda_{1N})$ 两种可能的情形:

(a) $\lambda_{0N} \geqslant 0$, $\lambda_{1N} \geqslant 0$, $\lambda_{0N}^2 + \lambda_{1N}^2 = 1$;

(b) $\lambda_{0N} \geqslant 0$, $\lambda_{1N} < 0$, $\lambda_{0N}^2 + \lambda_{1N}^2 = 1$.

在情形 (a) 中有

$$\delta_N = \lambda_{1N}\sqrt{2} + 3\lambda_{0N} - \lambda_{0N} \geqslant \sqrt{2}(\lambda_{1N} + \lambda_{0N}) \geqslant \sqrt{2},$$

而在情形 (b) 中有估计

$$\delta_N \geqslant |\lambda_{1N}| + \sqrt{2}\lambda_{0N} - \lambda_{0N} \geqslant (\sqrt{2}-1)(|\lambda_{1N}| + \lambda_{0N}) \geqslant \sqrt{2} - 1.$$

因此在所考虑的离散逼近问题序列中 AMP 不成立. △

由上面的讨论可看到, 例 6.61 中 AMP 不成立是由于, 在离散逼近过程中, 等式约束没有被扰动 (或被充分扰动), 同时费用泛函的最优值关于这样的扰动是不稳定的. 事实上, 任何控制 $u_N(t)$, 在某 $t \in T_N$ 等于 $(1,-2)$ 或 $(-\sqrt{2},-3)$, 并给出费用泛函的值 $J_N(u_N) < -1$, 对于约束 $x_{1N}(1)=0$ 是不可行的, 然而对该约束的适当扰动是可行的. 另一方面, U 中的这些点为 Hamilton-Pontryagin 函数提供了极大值. 这种情形在例 6.61 的离散时间系统中发生是由于在控制集 U 中无理数和有理数网点 T_N (对所有 $N \in \mathbb{N}$) 的不可公度性. 当然, 由实数的完备性, 这在连续时间系统中是不可能发生的.

6.4.5 在端点约束下的近似最大值原理: 证明及应用

经过上面所有讨论, 下面证明定理 6.59. 这里将证明分成三个主要步骤, 包括两个有独立意义的引理, 它们有助于对离散逼近隐含凸性的理解. 然后导出由一致上次可微函数描述的不等式约束问题 AMP 的上次微分扩展. 最后给出一些 AMP 在 (具有小步长的) 离散时间和连续时间系统中的典型应用.

对所有 $t \in T_N$, 当 $N \in \mathbb{N}$ 时, 设 $u_N(t) \in U$. 给定整数 r, 使得 $1 \leqslant r \leqslant N-1$, 定义控制 $u_N(\cdot)$ 的针形型变分如下. 考虑参数集 $\{\theta_j(N), v_j(N)\}_{j=1}^r$, 其中 $v_j(N) \in U$ 并且 $\theta_j(N)$ 为整数, 满足

$$0 \leqslant \theta_j(N) \leqslant N-1, \quad \text{并且} \quad \theta_j(N) \neq \theta_i(N) \quad (\text{如果} \quad j \neq i).$$

记 $\tau_{\theta_j(N)} := a + \theta_j(N) h_N$, 那么称

$$\widetilde{u}_N(t) := \begin{cases} v_j(N), & t = \tau_{\theta_j}(N), \\ u_N(t), & t \in T_N,\ t \neq \tau_{\theta_j(N)},\ j = 1, \cdots, r \end{cases} \tag{6.94}$$

为控制 $u_N(\cdot)$ 的具有参数 $\{\theta_j(N), v_j(N)\}$ 的 r- 针形变分. 当 $r=1$ 时, 控制 (6.94) 是 $u_N(\cdot)$ 的 (单) 针形变分, 但对 $r>1$, 它是 $u_N(\cdot)$ 的多针形变分. 这里引入的变分是离散时间的版本, 它对应于连续时间控制针形变分 (6.71) 和 (6.72), 然而它们有着本质的差别, 特别是在多针形的情形.

设 $\widetilde{x}_N(\cdot)$ 为有限差分系统

$$x_N(t+h_N) = x_N(t) + h_N f(x_N(t), u_N(t), t), \quad x_N(a) = x_0, \tag{6.95}$$

对应于控制变分 $\bar{u}_N(\cdot)$ 并且具有参数 $\{\theta_j(N), v_j(N)\}$ 的轨道; 下面在记号中通常略去其对 N 的依赖. 那么差 $\widetilde{x}_N(\cdot) - x_N(\cdot)$, 对 $r>1$, 记为 $\Delta^r_{\{\theta_j, v_j\}} x_N(\cdot)$; 对 $r=1$, 记为 $\Delta_{\theta,v} x_N(\cdot)$. 为方便起见, 它们分别称为多针形 (或 r- 针形) 和 (单) 针形轨道增量. 当 $t=b$ 时, 称为对应的端点增量.

下面首要的意图是建立当 $N\to\infty$ 时, 由参考控制 $u_N(\cdot)$ 的单针形变分生成的端点轨道增量的整数组合和一些多针形端点轨道增量之间的关系. 下面的结果本质上可视为这对连续系统至关重要的隐含凸性的近似有限差分版本.

设 $\{u_N(t)\}$ $(t\in T_N)$ 为参考控制序列, 并且对每个 $j=1,\cdots,p$, 设 $(\theta_j(N), v_j(N))$ 为 $u_N(\cdot)$ 的单针形变分的参量, 其中 p 是不依赖 N 的自然数. 当 $j=1,\cdots,p$ 时, 给定非负整数 m_j, 也不依赖 N, 考虑对应的针形轨道增量 $\Delta_{\theta_j,v_j}x_N(b)$, 为简单起见, 记为 $\Delta_{\theta,v,j}x(b)$. 对每个 $N=p,p+1,\cdots$, 构造 (单) 针形轨道增量的整数组合为

$$\Delta_N(p,m_j):=\sum_{j=1}^{p}m_j\Delta_{\theta,v,j}x_N(b).$$

证明它可表示为参考控制的一个多针形变分, 其中误差不超过 $o(h_N)$.

引理 6.62(针形轨道增量的整数组合) 设 $\{u_N(\cdot)\}$, $N\in\mathbb{N}$, 为一个恰当的参考控制序列, 设 $p\in\mathbb{N}$, $m_j\in\mathbb{N}\cup\{0\}$ $(j=1,\cdots,p)$ 不依赖 N, 并且设 $(\theta_j(N),v_j(N))$, $j=1,\cdots,p$, 为 (单) 针形变分的参数. 那么存在不依赖 N 的 $r\in\mathbb{N}$ 和 (6.94) 类型的 r- 针形变分参数 $\{\widetilde{\theta}_j(N),\widetilde{v}_j(N)\}_{j=1}^r$, 使得对相应的端点轨道增量有

$$\Delta_N(p,m_j)=\Delta^r_{\{\widetilde{\theta}_j,\widetilde{v}_j\}}x_N(b)+o(h_N)\quad(N\to\infty).$$

证明 首先建立由参考控制针形和多针形变分生成的端点轨道增量的简便表示, 这里并不要求构造恰当序列. 回顾上面矩阵乘法的记号, 记 $K>0$ 为 f 和 $\nabla_x f$ 沿着 $\{u_N(\cdot),x_N(\cdot)\}$ 的共同一致范数界限. 由 6.4.3 小节中的常驻假设, 该界限是存在的. 注意到, 沿着 (P_N) 的最优解的参数序列, 在下面的主要定理中实际上需要一致有界性, 但在该引理中没有用到.

先从参数 $(\theta(N),v(N))$ 生成的单针形变分开始. 由 (6.95) 和 f 关于 x 的光滑性即可得

$$\Delta_{\theta,v}x_N(\tau_i)=0,\quad i=0,\cdots,\theta,$$

$$\begin{aligned}\Delta_{\theta,v}x_N(\tau_{\theta+1})=&h_N[f(x_N(\tau_\theta),v,\tau_\theta)-f(x_N(\tau_\theta),u_N(\tau_\theta),\tau_\theta)]=:h_Ny,\\ \Delta_{\theta,v}x_N(\tau_{\theta+2})=&h_N[I+h_N\nabla_xf(x_N(\tau_{\theta+1}),u_N(\tau_{\theta+1}),\tau_{\theta+1})]y\\ &+h_No(\|\Delta_{\theta,v}x_N(\tau_{\theta+1})\|).\end{aligned}$$

那么由归纳法容易得到

$$\Delta_{\theta,v}x_N(b)=h_N\left\{\prod_{\theta+1}^{i=N-1}[I+h_N\nabla_xf(x_N(\tau_i),u_N(\tau_i),\tau_i)]\right\}y$$

$$
+h_N\sum_{k=\theta+2}^{N-1}\left\{\prod_{k}^{i=N-1}[I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i)]\right\}o(\|\Delta_{\theta,v}x_N(\tau_{k-1})\|)
$$
$$
+h_N o(\|\Delta_{\theta,v}x_N(\tau_{k-1})\|).
$$

由 (6.95) 和所作假设, 有 $\Delta_{\theta,v}x_N(t)=O(h_N)$ 对所有 $t\in T_N$ 关于 N 一致. 因此, 给定任何 $\varepsilon>0$, 存在 $N_\varepsilon\in\mathbb{N}$, 使得

$$
\|o(\|\Delta_{\theta,v}x_N(\tau_k)\|)\|\leqslant\varepsilon h_N,\quad k=\theta+2,\cdots,N-1,\ N\geqslant N_\varepsilon,
$$

这蕴涵估计

$$
\left\|\sum_{k=\theta+2}^{N-1}\left\{\prod_{k}^{i=N-1}[I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i)]\right\}o(\|\Delta_{\theta,v}x_N(\tau_{k-1})\|)\right\|
$$
$$
\leqslant\varepsilon h_N\sum_{k=\theta+2}^{N-1}\prod_{k}^{i=N-1}(1+h_N K)^i\leqslant\frac{\varepsilon}{K}\exp(K(b-a)).
$$

将这与上面 $\Delta_{\theta,v}x_N(b)$ 的公式结合起来, 可得到由参考控制单针形变分生成的端点轨道增量的有效表示

$$
\begin{aligned}
\Delta_{\theta,v}x_N(b)=h_N&\left\{\prod_{\theta+1}^{i=N-1}[I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i)]\right\}y\\
&+o(h_N)\quad(N\to\infty),
\end{aligned}
\tag{6.96}
$$

其中当 $N\to\infty$ 时, $o(h_N)/h_N\to 0$ 不依赖针形参数 $\theta=\theta(N)$ 和 $v=v(N)$.

现在考虑由多针形变分 (6.74) 生成的具有参数 $\{\theta_j(N),v_j(N)\}_{j=1}^r$ 的端点轨道增量. 类似 (6.96), 可导出下面的表达式

$$
\begin{aligned}
\Delta^r_{\{\theta_j,v_j\}}x_N(b)=h_N&\left\{\sum_{j=1}^{r}\left[\prod_{\theta_j+1}^{i=N-1}(I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i))\right]y_j\right\}\\
&+o(h_N)\quad(N\to\infty),
\end{aligned}
\tag{6.97}
$$

其中 $o(h_N)$ 不依赖 $\{\theta_j(N),v_j(N)\}$, 但依赖于 r 多个变动点, 并且

$$
y_j:=f(x_N(\tau_{\theta_j}),v_j,\tau_{\theta_j})-f(x_N(\tau_{\theta_j}),u_N(\tau_{\theta_j}),\tau_{\theta_j}),\quad j=1,\cdots,r.
$$

下面假设控制序列 $\{u_N(\cdot)\}$ 是恰当的, 并且证明该引理中阐述的主要关系. 不失一般性, 假设网点

$$
\tau_{\theta_j(N)}:=a+\theta_j(N)h_N,\quad j=1,\cdots,p,
$$

当 $N\to\infty$ 时, 收敛于某个数 $\bar{\tau}_j\in[a,b]$, $j=1,\cdots,p$. 首先验证

$$\bar{\tau}_i\neq\bar{\tau}_j \quad (i\neq j), \quad \bar{\tau}_j\neq b \quad (i,j\in\{1,\cdots,p\}) \tag{6.98}$$

的情形. 对每个 $N\geqslant p$, 给定整数组合 $\Delta_N(p,m_j)$ 的参数, 然后取不依赖 N 的数 $r:=m_1+\cdots+m_p$, 考虑由多针形变分 $\widetilde{u}_N(t)$ 生成的端点轨道增量 $\Delta^r_{\{\widetilde{\theta}_{jq},\widetilde{v}_{jq}\}}x_N(b)$, 其中

$$\widetilde{u}_N(t):=\begin{cases} v_j(N), & \text{如果 } t=\tau_{\theta_j+q}(N), \\ u_N(t), & \text{如果 } t\neq\tau_{\theta_j+q}(N),\ t\in T_N, \end{cases} \tag{6.99}$$

只要 $j=1,\cdots,p$ 和 $q=0,\cdots,m_j-1$, 并且

$$\widetilde{\theta}_{jq}(N):=\theta_j(N)+q, \quad \widetilde{v}_{jq}(N):=v_j(N) \quad (\forall j,q).$$

由假设 (6.98), 这些针形控制变分对所有大的 N 是有良好定义的. 利用对应于端点增量的表达式 (6.97), 有

$$\Delta^r_{\{\widetilde{\theta}_{iq},\widetilde{v}_{jq}\}}x_N(b)=h_N\left\{\sum_{j=1}^{p}\sum_{q=1}^{m_j}\left[\prod_{\theta_j+q}^{i=N-1}(I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i))\right]y_{jq-1}\right\}$$
$$+o(h_N) \quad (N\to\infty),$$

其中 $o(h_N)$ 有一致估计并且

$$y_{jq}:=f(x_N(\tau_{\theta_j+q}),v_j,\tau_{\theta_j+q})-f(x_N(\tau_{\theta_j+q}),u_N(\tau_{\theta_j+q}),\tau_{\theta_j+q}).$$

由 $\{u_N(\cdot)\}$ 的恰当性和 f 关于所有变量的连续性, 当 $N\to\infty$ 时, 有 $y_{j\theta}-y_{j0}\to 0$, 这蕴涵表达式

$$\Delta^r_{\{\widetilde{\theta}_{iq},\widetilde{v}_{jq}\}}x_N(b)=h_N\left\{\sum_{j=1}^{p}m_j\prod_{\theta_j+1}^{i=N-1}[I+h_N\nabla_x f(x_N(\tau_i),u_N(\tau_i),\tau_i)]\,y_j\right\}$$
$$+o(h_N) \quad (N\to\infty),$$

其中 y_j 在 (6.97) 中定义. 比较这个表达式和由具有参数 $(\theta_j(N),v_j(N))$, $j=1,\cdots,p$ 的单针形变分生成的端点轨道增量公式 (6.96), 并考虑到 $\Delta_N(p,m_j)$ 的表达式, 可得到在上面关于极限点 $\bar{\tau}_j$ 的要求 (6.98) 下引理中的结论.

现在假设这些要求不满足, 那么只需研究下面两种极端情形:

(a) $\bar{\tau}_1=\bar{\tau}_2=\cdots=\bar{\tau}_p\neq b$,

(b) $\bar{\tau}_1=\bar{\tau}_2=\cdots=\bar{\tau}_p=b$.

这与 (6.98) 一起涵盖了 $[a,b]$ 中极限点 $\bar{\tau}_j$ 的所有可能位置. 下面在 (a) 和 (b) 这两种情形中, 给出多针形变分 (6.99) 相应的修改, 类似上面的讨论, 就可导出引理的结论.

在情形 (a) 中, 重新排列 $(\theta_j(N), v_j(N))$, $j=1,\cdots,p$, 使得 $\theta_1<\cdots<\theta_p$ (不失一般性, 假设所有的 θ_j 都不相同). 为了简便起见, 指标 θ_j 与相应的网点 τ_{θ_j} 等同. 那么像 (6.99) 那样在点 $\theta_1, \theta_1+1, \cdots, \theta_1+m_1-1$ 构造 $u_N(\cdot)$ 的变分. 假设对应于参数 (θ, v_i), $1\leqslant i\leqslant p-1$, 的控制变分已构造好, 然后构造 (θ_{i+1}, v_{i+1}). 记 θ_0 为 $\{\theta_j\}$ 中最大的点, 其中在这些点已经构造好控制变分. 如果 $\theta_0<\theta_{i+1}$, 那么像 (6.99) 那样, 在点 $\theta_{i+1}, \theta_{i+1}+1, \cdots, \theta_{i+1}+m_{i+1}$ 构造 $u_N(\cdot)$ 的变分. 如果 $\theta_0\geqslant\theta_{i+1}$, 那么在点 $\theta_0+1, \cdots, \theta_0+m_{i+1}$ 构造相同类型的变分. 可以验证, 在情形 (a) 中, 这样构造的多针形变分确保满足引理的结论.

在情形 (b) 中, 重新排列 $(\theta_j(N), v_j(N))$, $j=1,\cdots,p$, 使得 $\theta_1>\theta_2>\cdots>\theta_p$, 然后对称于情形 (a), 即从右到左, 构造对应于 $u_N(\cdot)$ 的多针形变分. 这样就完成了引理的证明. △

下一个结果给出了类似引理 6.44 的序列有限差分, 并且可认为是离散逼近问题中隐含凸性的某种近似 (不是确切/极限) 的表示, 它没有用到抽象的时间连续性. 为此, 在离散逼近过程中需要区分本性和非本性约束, 这在下面是重要的.

定义 6.63(有限差分系统的本性和非本性不等式约束) 不等式端点约束

$$\varphi_i(x_N(b))\leqslant\gamma_{iN}, \quad 对某 \quad i\in\{1,\cdots,m\},$$

对问题 (P_N) 的可行解序列 $\{u_N(\cdot), x_N(\cdot)\}$, 沿着自然数序列 $\mathcal{M}\subset\mathbb{N}$, 是本性的. 如果

$$\varphi_i(x_N(b))-\gamma_{iN}=O(h_N) \quad (h_N\to\infty),$$

即存在实数 $K_i\geqslant 0$ 满足

$$-K_i h_N\leqslant\varphi_i(x_N(b))-\gamma_{iN}\leqslant 0 \quad (N\to\infty,\ N\in\mathcal{M}).$$

该约束对序列 $\{u_N(\cdot), x_N(\cdot)\}$ 沿着 $\mathcal{M}$ 是非本性的, 如果只要 $K>0$, 那么存在 $N_0\in\mathbb{N}$, 使得

$$\varphi_i(x_N(b))-\gamma_{iN}\leqslant -Kh_N, \quad \forall N\geqslant N_0, \quad N\in\mathcal{M}.$$

离散逼近序列中本性约束的概念对应于非参数最优化问题起作用约束的概念. 不失一般性, 假设对参数问题 (P_N) 的最优解序列 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 前 $l\in\{1,\cdots,m\}$ 个不等式约束是本性的, 但其他 $m-l$ 个约束沿着所有自然数, 即 $\mathcal{M}=\mathbb{N}$, 是非本性的.

给定问题 (P_N), $N \in \mathbb{N}$ 的最优解 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 构造线性化映像集

$$S_N := \{(y_0, \cdots, y_l) \in \mathbb{R}^{l+1} | y_i = \langle \nabla\varphi_i(\bar{x}_N(b)), \Delta_{\theta,v}\bar{x}_N(b)\rangle\}. \tag{6.100}$$

它是由涉及费用函数、本性不等式约束的梯度以及对应于最优控制的所有单针形变分的端点轨道增量生成的. 随后的目的是要证明集合 (6.100) 的凸包序列 $\{\mathrm{co}\, S_N\}$, 当 $h_N \to 0$ 时, 可以平移某数量级为 $o(h_N)$ 的量, 使得所得到的集合与 $\mathbb{R}^{l+1}$ 中禁止点的凸集 $\mathbb{R}^{l+1}_<$ 不相交, 其中

$$\mathbb{R}^{l+1}_< := \{(y_0, \cdots, y_l) | y_i < 0 \text{ 对所有 } i = 0, \cdots, l\}.$$

引理 6.64(具有不等式约束的离散逼近问题的隐含凸性和本原最优性条件) 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (P_N) 的最优解序列, $\varphi_i = 0$, $i = m+1, \cdots, m+r$ (没有扰动的等式约束). 除了所作的常驻假设外, 假设端点函数 φ_i 对所有 $i = 0, \cdots, m$ 在序列 $\{\bar{x}_N(\cdot)\}$ 的极限点附近是连续可微的. 也假设控制序列 $\{\bar{u}_N(\cdot)\}$ 是恰当的, 且对 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 前 $l \in \{1, \cdots, m\}$ 个不等式约束是本性的, 但其他的是非本性的. 那么存在一个 $(l+1)$- 维数量级为 $o(h_N)$ (当 $h_N \to 0$ 时) 的量, 使得

$$(\mathrm{co}\, S_N + o(h_N)) \cap \mathbb{R}^{l+1}_< = \varnothing, \quad \text{对所有大的 } N \in \mathbb{N} \text{ 成立}. \tag{6.101}$$

证明 对每个 N 和不依赖 N 的固定 $r \in \mathbb{N}$, 考虑由最优控制 $\bar{u}_N(\cdot)$ 的多针形变分生成的端点轨道增量, 其中 $\{\theta_j(N), v_j(N)\}_{j=1}^r$ 是 (6.94) 中的变分参数. 构造一列向量

$$y_N = (y_{N0}, \cdots, y_{Nl}) \in \mathbb{R}^{l+1}, \quad \text{其中} \quad y_{Ni} := \langle \nabla\varphi_i(\bar{x}_N(b)), \Delta^r_{\{\theta_j, v_j\}}\bar{x}_N(b)\rangle,$$

证明存在一个 $(l+1)$- 维数量级为 $o(h_N)$ (当 $N \to \infty$ 时) 的序列, 使得

$$y_N + o(h_N) \notin \mathbb{R}^{l+1}_< \quad (N \to \infty). \tag{6.102}$$

事实上, 由表示 (6.97) 和所作的假设有

$$\|\Delta^r_{\{\theta_j, v_j\}}\bar{x}_N(b)\| \leqslant \mu h_N \quad (\forall t \in T_N, N \in \mathbb{N}),$$

其中 $\mu > 0$ 依赖于 r 但不依赖 $\{\theta_j(N), v_j(N)\}_{j=1}^r$. 由 $\bar{x}_N(\cdot)$ 在无等式约束扰动的问题 (P_N) 中的最优性, 对每个 $N \in \mathbb{N}$, 存在指标 $i_0(N) \in \{0, \cdots, m\}$ 使得

$$\varphi_{i_0}(\bar{x}_N(b) + \Delta^r_{\{\theta_j, v_j\}}\bar{x}_N(b)) - \varphi_{i_0}(\bar{x}_N(b)) \geqslant 0.$$

因为对 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 只有前 l 个不等式约束是本性的, 只要 N 充分大, 该不等式对某 $i_0 \in \{0, \cdots, l\}$ 成立. 考虑数

$$\delta_N := \max_{0 \leqslant i \leqslant l} \sup\left\{ |\varphi_i(\bar{x}_N(b) + \Delta x) - \varphi_i(\bar{x}_N(b)) - \langle \nabla\varphi_i(\bar{x}_N(b)), \Delta x\rangle| \;\middle|\; \|\Delta x\| \leqslant \mu h_N \right\}.$$

由 φ_i 的光滑性假设, 当 $N \to \infty$ 时, $\delta_N / h_N \to 0$ 关于变分是一致的. 这意味着

$$y_{Ni_0} + \delta_N \geqslant 0 \quad (N \to \infty).$$

这就证明了 (6.102), 并且数量 $o(h_N) := (0, \cdots, \delta_N, \cdots, 0) \in \mathbb{R}^{l+1}$, 其中 δ_N 出现在第 $i_0(N)$ 个位置上.

下一个目标是, 对由最优控制单针形变分生成的端点轨道增量凸组合, 得到类似 (6.102) 的估计. 在这样的整数组合情形下, 由前面的引理 6.62, 直接由该估计可得到对应的 (6.102). 现在证明, 凸组合情形实际上可约化为整数组合的情形.

考虑参数列 $(\theta_j(N), v_j(N))$, $j = 1, \cdots, p$, 它生成了最优控制 $\bar{u}_N(\cdot)$ 的单针形变分并且 $p \in \mathbb{N}$, 然后定义凸组合

$$\begin{aligned} y_{Ni}(p, \alpha) &:= \sum_{j=1}^{p} \alpha_j(N) \left\langle \nabla \varphi_i(\bar{x}_N(b)), \Delta_{\theta,v,j} \bar{x}_N(b) \right\rangle \\ &(\text{当 } \alpha_j(N) \geqslant 0,\ \alpha_1(N) + \cdots + \alpha_p(N) = 1,\ i = 0, \cdots, l \text{ 时}). \end{aligned} \tag{6.103}$$

在上面的组合中固定 (p, α), 并取 $y_N(p, \alpha) \in \mathbb{R}^{l+1}$, 其分量为 $y_{Ni}(p, \alpha)$, 假设存在数 $N_0 \in \mathbb{N}$, 使得

$$y_N(p, \alpha) \in \mathbb{R}^{l+1}_{<} \quad (N \geqslant N_0).$$

现在证明对每个自然数 $N \geqslant N_0$, 存在指标 $i_0 = i_0(N) \in \{0, \cdots, l\}$, 使得

$$0 > y_{Ni_0}(p, \alpha) = o(h_N) \quad (h_N \to \infty). \tag{6.104}$$

如若不然, 可找到子列 $\mathcal{M} \subset \mathbb{N}$, 使得

$$\lim_{N \to \infty} \frac{y_{Ni}(p, \alpha)}{h_N} := \beta_i < 0 \quad (\text{当 } N \in \mathcal{M} \text{ 时, 对所有 } i = 0, \cdots, l).$$

不失一般性, 假设 $\mathcal{M} = \{p, p+1, \cdots\}$, $\beta_i > -\infty$ 并且对每个 $j = 1, \cdots, p$, 当 $N \to \infty$ 时, 序列 $\{\alpha_j(N)\}$ 收敛于某 $\alpha_j^0 \in \mathbb{R}$. 给定 $\nu > 0$, 定义 p 个整数 k_j 为

$$k_j = k_j(\nu) := \left[\frac{\alpha_j^0}{\nu} \right] \quad (j = 1, \cdots, p).$$

构造整数组合 $y_{N_i}(p, k)$ $(i = 0, \cdots, l)$ 为

$$y_{Ni}(p, k) := \frac{y_{Ni}(p, \alpha^0)}{\nu} + \sum_{j=1}^{p} \left(k_j - \frac{\alpha_j^0}{\nu} \right) \left\langle \nabla \varphi_i(\bar{x}_N(b)), \Delta_{\theta,v,j} \bar{x}_N(b) \right\rangle,$$

其中 $k := (k_1, \cdots, k_p)$, $\alpha^0 := (\alpha_1^0, \cdots, \alpha_l^0)$.

设 $\mu>0$ 为在证明 (6.102) 式中选取的常数 (其中 $r=1$), 且设 $\kappa>0$ 为 $\varphi_i(\bar{x}_N(b))$ 和 $\nabla\varphi_i(\bar{x}_N(b))$ $(i=0,\cdots,l)$ 的一致范数界限. 选取 $i_1\in\{0,\cdots,l\}$, 定义 $\nu>0$, 使得

$$|\beta_{i_1}|=\min_{0\leqslant i\leqslant l}|\beta_i|,\quad \nu:=\frac{\beta_{i_1}}{\beta_{i_1}-p\kappa\mu}.$$

那么有估计

$$\lim_{N\to\infty}\frac{y_{Ni}(p,k)}{h_N}\leqslant\beta_i-\frac{\beta_i\mu\kappa p}{\beta_{i_1}}+\mu\kappa p\leqslant\beta_i<0\quad(i=0,\cdots,l).$$

由引理 6.62 关于 (单) 控制变分生成的端点轨道增量的整数组合表示, 显然和 (6.102) 矛盾. 这就证明了 (6.104).

最后, 证明所需的关系式 (6.101). 对所有大的 $N\in\mathbb{N}$, 当 $\operatorname{co}S_N\cap\mathbb{R}^{l+1}_<=\varnothing$ 时, 无需证明. 假设沿着子列 $\{N\}$ (不失一般性, 取为整个自然数集 $\mathbb{N}$), $\operatorname{co}S_N\cap\mathbb{R}^{l+1}_<\neq\varnothing$. 对每个 $N\in\mathbb{N}$, 定义

$$\sigma_N:=-\inf\left\{\max_{0\leqslant i\leqslant l}y_i\,\middle|\,y=(y_0,\cdots,y_l)\in\operatorname{co}S_N\cap\mathbb{R}^{l+1}_<\right\},$$

其中, 在所作的假设下, 最小值在某 $y_N\in\mathbb{R}^{l+1}_<$ 处达到. 利用经典的 Carathéodory定理, 将 y_N 表示为凸组合形式 (6.103), 其中 $p=l+2$. 现在利用 (6.104), 可找到指标 $i_0=i_0(N)$, 使得

$$\sigma_N=-\max\{y_{Ni}|i=0,\cdots,l\}\leqslant-y_{Ni_0}=o(h_N)\quad(N\to\infty),$$

这蕴涵着 (6.101) 式并且具有 $(l+1)$- 维平移 $o(h_N):=(\sigma_N,\cdots,\sigma_N)$. 因此完成了引理的证明. △

定理 6.59 的证明 现在已有了所有主要工具来完成定理的证明. 从问题 (P_N) 中只有扰动的不等式约束开始, 即 $\varphi_i=0,\ i=m+1,\cdots,m+r$. 因为, 不失一般性, 假设前 $l\leqslant m$ 个不等式约束对最优解序列 $\{\bar{u}_N(\cdot),\bar{x}_N(\cdot)\}$ 是本性的, 而其余的 $m-l$ 个不等式约束对该序列是非本性的, 所以由定义 6.63 得到

$$\varphi_i(\bar{x}_N(b))-\gamma_{iN}=O(h_N)\quad(N\to\infty,\quad i=1,\cdots,l).$$

在 (6.101) 式中的凸集, 利用引理 6.64 和经典的凸集分离定理, 可找到单位向量序列 $(\lambda_{0N},\cdots,\lambda_{lN})\in\mathbb{R}^{l+1}$ 将这些集合分离开. 考虑到 (6.101) 式中集合的结构, 对具有参数 $(\theta(N),v(N))$ 最优控制的任何 (单) 针形变分, 易得

$$\lambda_{iN}\geqslant0\quad(i=0,\cdots,l),\quad\lambda_{0N}^2+\cdots+\lambda_{lN}^2=1,$$

$$\sum_{i=0}^{l}\lambda_{iN}\left\langle\nabla\varphi_i(\bar{x}_N(b)),\Delta_{\theta,v}\bar{x}_N(b)\right\rangle+o(h_N)\geqslant0\quad(N\to\infty).$$

现在记

$$\lambda_{iN} := 0 \quad (\text{对 } i = l+1, \cdots, m, \text{ 当 } N \to \infty \text{ 时}).$$

类似关于自由端点问题定理 6.50 的证明, 当 N 充分大时, 对所有 $v \in U$ 和 $t \in T_N$ 有

$$h_N \Big[H(\bar{x}_N(t), p_N(t+h_N), v, t) - H(\bar{x}_N(t), p_N(t+h_N), \bar{u}_N(t), t) \Big] + o(h_N) \leqslant 0.$$

其中每个 $p_N(\cdot)$ 满足具有横截性条件 (6.93) 的伴随系统 (6.86), 并且对所考虑的不等式约束问题 (P_N), $\lambda_{0N}, \cdots, \lambda_{mN}$ 显然满足条件 (6.91) 和 (6.92). 如定理 6.50 的证明那样, 由反证法, 上面的 Hamilton 不等式直接蕴涵着近似最大值条件 (6.85). 这就完成了具有不等式约束问题 (P_N) 情形定理的证明.

现在考虑 (P_N) 的一般情形, 即具有扰动的等式约束的情形. 每一个约束 $|\varphi_{iN}(x_N(b))| \leqslant \xi_{iN}$ $(i = m+1, \cdots, m+r)$ 明显地可分为两个不等式约束

$$\varphi_{iN}^+(x_N(b)) := \varphi_i(x_N(b)) - \xi_{iN} \leqslant 0,$$
$$\varphi_{iN}^-(x_N(b)) := -\varphi_i(x_N(b)) - \xi_{iN} \leqslant 0.$$

接下来证明, 如果其中的一个约束对 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, 沿着某子列 $\mathcal{M} \subset \mathbb{N}$ 是本性的, 那么在相容性条件 (6.80) 下, 另一个约束沿着相同的子列也是非本性的. 事实上, 为了明确起见, 假设约束 $\varphi_{iN}^+(\bar{x}_N(b)) \leqslant 0$ 对某 $i \in \{m+1, \cdots, m+r\}$ 沿着 $\mathcal{M}$ 是本性的. 那么由 (6.80), 当 $N \in \mathcal{M}$ 时, 对任何 $K > 0$ 有

$$\varphi_{iN}^-(\bar{x}_N(b)) = -\varphi_i(\bar{x}_N(b)) + \xi_{iN} - 2\xi_{iN} = -\varphi_{iN}^+(\bar{x}_N(b)) - 2\xi_{iN} \leqslant Kh_N.$$

这意味着约束 $\varphi_{iN}^-(\bar{x}_N(b)) \leqslant 0$ 是非本性的. 这样, 利用上面证明的该定理的不等式情形, 可找到乘子 λ_{iN}^+ 和 λ_{iN}^-, 满足

$$\lambda_{iN}^+ \cdot \lambda_{iN}^- = 0 \quad (\text{对 } i = m+1, \cdots, m+r, \text{ 当 } N \to \infty \text{ 时}).$$

最后记

$$\lambda_{iN} := \lambda_{iN}^+ - \lambda_{iN}^-, \quad i = m+1, \cdots, m+r,$$

完成了定理的证明. △

注 6.65(在轨道的两个端点和中间点有约束控制问题的 AMP)　上面发展的方法可用来导出更一般的 (P_N) 类型的离散逼近问题的 AMP 形式的必要最优性条件, 这类问题具有费用函数 $\varphi_0(x_N(a), x_N(b))$ 和在可行轨道的两个端点上都有的约束

$$\varphi_i(x_N(a), x_N(b)) \leqslant \gamma_{iN}, \quad i = 1, \cdots, m,$$
$$|\varphi_i(x_N(a), x_N(b))| \leqslant \xi_{iN}, \quad i = m+1, \cdots, m+r.$$

在与定理 6.50 和定理 6.59 中关于初始数据相同的假设下 AMP 成立, 且对在最优轨道左端点具有额外近似横截性条件

$$p_N(a) = \sum_{i=0}^{m+r} \lambda_{iN} \nabla_{x_a} \varphi_i(\bar{x}_N(a), \bar{x}_N(b)),$$

其中 $\nabla_{x_a}\varphi_i$ 表示函数 $\varphi_i(x_a, x_b)$ 在最优点的偏导数.

对于具有目标 $\varphi_0 = \varphi(x_a, x_\tau, x_b)$ 和中间状态约束类型

$$\varphi_i(x_N(a), x_N(\tau_N), x_N(b)) \leqslant \gamma_{iN}, \quad i = 1, \cdots, m,$$

$$|\varphi_i(x_N(a), x_N(\tau_N), x_N(b))| \leqslant \xi_{iN}, \quad i = m+1, \cdots, m+r$$

的 (P_N) 的问题, 可导出类似的结果, 其中 $\tau_N \in T_N$ 是网格的中间点. 这类问题的 AMP 涉及额外的跳跃型确切条件:

$$\begin{aligned} p_N(\tau_N + h_N) - p_N(\tau_N) = & \sum_{i=0}^{m+r} \lambda_{iN} \nabla_{x_\tau} \varphi_i(\bar{x}_N(a), \bar{x}_N(\tau_N), \bar{x}_N(b)) \\ & - h_N \nabla_x H(\bar{x}_N(\tau_N), p_N(\tau_N + h_N), \bar{u}_N(\tau_N), \tau_N). \end{aligned}$$

注意到, 在该情形中, 要求伴随系统 (6.86) 对 $p_N(\cdot)$ 在点 $t \in T_N \setminus \tau_N$ 成立.

下面给出定理 6.59 在非光滑问题 (P_N) 中的扩展, 其中费用和不等式约束函数 φ_i, $i = 0, \cdots, m$ 是一致上次可微的. 在此情形中, 横截性条件以上次微分的形式得到.

定理 6.66(具有上次微分横截性条件的约束非光滑问题的 AMP) 在定理 6.59 除了 $\varphi_i, i = 0, \cdots, m$ 的光滑性外的所有假设下, 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (P_N) 对 $N \in \mathbb{N}$ 的最优解. 假设函数 φ_i $(i = 0, \cdots, m)$ 在 $\{\bar{x}_N(b)\}$ 的极限点附近是一致上次可微的. 那么对任何上次梯度序列 $x_{iN}^* \in \widehat{\partial}^+ \varphi_i(\bar{x}_N(b))$, $i = 0, \cdots, m$, 存在 $\{\lambda_{iN} | i = 0, \cdots, m+r\}$, 使得所有条件 (6.85), (6.86), (6.91) 和 (6.92) 都成立, 并且

$$p_N(b) = -\sum_{i=0}^{m} \lambda_{iN} x_{iN}^* - \sum_{i=m+1}^{m+r} \lambda_{iN} \nabla \varphi_i(\bar{x}_N(b)).$$

证明 对 $i = 0, \cdots, m$ 和 $N \in \mathbb{N}$, 给定 $x_{iN}^* \in \widehat{\partial}^+ \varphi_i(\bar{x}_N(b))$, 然后构造 (6.100) 中 S_N 的非光滑版本为

$$S_N := \{(y_0, \cdots, y_l) \in \mathbb{R}^{l+1} | y_i = \langle x_{iN}^*, \Delta_{\theta,v} \bar{x}_N(b) \rangle\},$$

那么可得到类似引理 6.64 的结果和证明. 唯一不同的是现在得到的同样结果是基于不等式

$$\varphi_i(\bar{x}_N(b) + \Delta x) - \varphi_i(\bar{x}_N(b)) - \langle x_{iN}^*, \Delta x \rangle + o(\|\Delta x\|) \leqslant 0$$

(这是由于 φ_i, $i=1,\cdots,l$ 的一致上次可微性), 而不是光滑情形中引理 6.64 证明中用到的等式

$$\varphi_i(\bar{x}_N(b)+\Delta x)-\varphi_i(\bar{x}_N(b))-\langle\nabla\varphi_i(\bar{x}_N(b)),\Delta x\rangle+o(\|\Delta x\|)=0.$$

将分离定理用到上面的凸集中有

$$\sum_{i=0}^{l}\langle x_{iN}^*,\Delta_{\theta,v}\bar{x}_N(b)\rangle+o(h_N)\geqslant 0.$$

类似定理 6.59 的证明, 就得到具有上次微分横截性条件的近似最大值原理. △

注 6.67 (通过离散逼近得到的连续时间系统的次最优性条件) 上面得到的满足离散逼近问题的 AMP 结果, 可以用来导出连续时间系统的以某种 ε- 最大值原理形式表示的次最优性条件. 在 5.1.4 小节中已讨论过, 特别是在无穷维空间的框架下, 次最优性条件对最优化理论和应用的重要性. 5.1.4 小节中的结果和讨论主要涉及泛函约束的数学规划问题. 在连续时间系统的最优控制中 (即使对有限维状态空间), 众所周知, 在具有非凸速率的系统中, 最优解经常不存在, 因此次最优性条件成为主要考虑对象. 在这类情形中, "几乎最优" (次最优) 解的 "几乎必要条件" 为最优化问题提供了重要的信息, 这从定性和定量/数值的角度来看都是很关键的.

由上面关于离散逼近的值稳定性 (参见 6.1.4 小节中的定理 6.14) 有, 给定任何 $\varepsilon>0$, 考虑的离散逼近问题 (P_N) 的最优解 $\{\bar{u}_N(\cdot),\bar{x}_N(\cdot)\}$ 可用来构造相应的连续时间问题 (P) 的 ε- 最优解 $\{u_\varepsilon(\cdot),x_\varepsilon(\cdot)\}$, 满足

$$\varphi_0(x_\varepsilon(b))\leqslant\inf J[x,u]+\varepsilon,$$

$$\varphi_i(x_\varepsilon(b))\leqslant\varepsilon\quad(i=1,\cdots,m),\quad|\varphi_i(x_\varepsilon(b))|\leqslant\varepsilon\quad(i=m+1,\cdots,m+r).$$

此外, 连续时间问题 (P) 的 ε- 最优控制总可取为 $[a,b]$ 上的分段常数.

现在利用定理 6.59 中提供的 AMP 形式的离散逼近问题 (P_N) 的必要最优性条件, 可得到下面关于问题 (P) 的次最优解的 ε- 最大值原理: 存在乘子 $(\lambda_0,\cdots,\lambda_{m+r})\in\mathbb{R}^{m+r+1}$, 满足

$$\lambda_i\geqslant 0\quad(i=0,\cdots,m),\quad\lambda_0^2+\cdots+\lambda_{m+r}^2=1,$$

$$|\lambda_i\varphi_i(x_\varepsilon(b))|\leqslant\varepsilon\quad(i=1,\cdots,m),$$

使得只要 $u\in U$ 和 $t\in[a,b]$, 就有

$$H(x_\varepsilon(t),p_\varepsilon(t),u_\varepsilon(t),t)\geqslant H(x_\varepsilon(t),p_\varepsilon(t),u,t)-\varepsilon,$$

其中 $p_\varepsilon(\cdot)$ 是伴随系统

$$\dot{p} = -\nabla H(x_\varepsilon(t), p, u_\varepsilon(t), t), \quad t \in [a,b]$$

对应的轨道, 并且具有横截性条件

$$p_\varepsilon(b) = -\sum_{i=0}^{m+r} \nabla\varphi_i(x_\varepsilon(b)).$$

对于在某些点 $\tau_j \in (a,b)$ 具有中间状态约束的连续时间问题以及在 $t=a$ 和 $t=b$ 具有端点约束的问题, 类似的结果也成立; 请比照注 6.65. 在端点约束的情形, 可得在点 $t=a$ 的 ε- 横截性条件

$$\left| p_\varepsilon(a) - \sum_{i=0}^{m+r} \lambda_i \nabla_{x_a}\varphi_i(x_\varepsilon(a), x_\varepsilon(b)) \right| \leqslant \varepsilon.$$

然而, 注意到定理 6.66 中上次微分形式的 AMP, 对连续时间系统, 不适合用来导出类似的次最优性结果, 这是因为 Fréchet上次微分 $\widehat{\partial}^+\varphi(\cdot)$ 对非光滑函数一般没有所要的连续性性质.

本小节的最后举例说明 AMP 在最优化具有小步长离散化的约束离散时间系统中的应用. 首先, 从定理 6.50 (和定理 6.59) 的证明中观察到, 当 $N\to\infty$ 时, 费用和约束函数在问题 (P_N) 的最优控制 $\bar{u}_N(\cdot)$, 与在 $u \in U$ 上最大化 Hamilton-Pontryagin 函数 $H(\bar{x}_N(t), p_N(t), \cdot, t)$ 的控制 $u_N(\cdot)$ 之间的差值, 是 $o(h_N)$ 数量级的. 事实上, 这意味着近似最大值原理在具有小步长 h_N 的离散时间系统最优化中的应用实际上有着与确切情形, 即离散最大值原理, 同样的效果. 有鉴于此, 现在利用 AMP 求解一些化学过程中出现的离散逼近问题.

例 6.68(AMP 在催化置换最优化中的应用) 考虑下面的二维连续时间系统的最优控制问题 (P), 它出现在催化置换模型中; 参见例如, Fan 和 Wang [426]:

$$\begin{cases} \min & J[u,x] = \varphi_0(x(1)) := x_1(1) \\ \text{s.t.} & \dot{x}_1 = -u_1(u_1+u_2), \quad \dot{x}_2 = u_1, \quad x_1(0)=x_2(0)=0, \quad t \in T := [0,1], \\ & u(t) = (u_1(t), u_2(t)) \in U := \{(u_1,u_2)\in\mathbb{R}^2 | 0 \leqslant u_1, u_2 \leqslant 2\}, \\ & \varphi_1(x(1)) := x_2(1) \leqslant 0. \end{cases}$$

为了数值地求解该问题, 构造一列离散逼近问题 (P_N):

$$\begin{cases} \min & J_N[u_N, x_N] := \varphi_0(x_N(1)) := x_{1N}(1) \\ \text{s.t.} & x_{1N}(t+h_N) = x_{1N}(t) - h_N u_{1N}[u_{1N}(t) + u_{2N}(t)], \quad x_{1N}(0) = 0, \\ & x_{2N}(t+h_N) = x_{2N}(t) + h_N u_{1N}(t), \quad x_{2N}(0) = 0, \quad h_N := N^{-1}, \\ & 0 \leqslant u_{1N}(t) \leqslant 2, \quad 0 \leqslant u_{2N}(t) \leqslant 2, \quad t \in T_N := \{0, h_N, \cdots, 1-h_N\}, \\ & \varphi_1(x_N(1)) = x_{2N}(1) \leqslant 0 \quad (N\to\infty). \end{cases}$$

因为 (P_N) 中的"容许速率"集 $f(x,U,t)$ 不是凸的, 所以 (确切) 离散最大值原理不能用来寻找这些问题的最优控制. 这里利用定理 6.59 证明过的近似最大值原理.

对每个 $N \in \mathbb{N}$, 具有横截性条件 (6.93) 的伴随系统 (6.86) 的对应轨道 $p_N(t) = (p_{1N}(t), p_{2N}(t))$ 为

$$p_{1N}(t) = -\lambda_{0N}, \quad p_{2N}(t) = -\lambda_{1N} \quad (t \in T_N),$$

并且沿着该轨道, Hamilton-Pontryagin 函数为

$$H_N(u,t) = u_1(\lambda_{0N}u_1 + \lambda_{0N}u_2 - \lambda_{1N}), \quad t \in T_N.$$

接下来确定在控制区域 U 上极大化 Hamilton-Pontryagin 函数的控制 $\widehat{u}_N(t) = (\widehat{u}_{1N}(t), \widehat{u}_{2N}(t))$. 由 (6.92) 中的规范化条件, 容易看出, 当 $\lambda \in (0,1)$ 时, 这类控制在 $(u_1,u_2) \in U$ 上极大化函数

$$H_\lambda(u_1,u_2) := u_1(\lambda u_1 + \lambda u_2 - \sqrt{1-\lambda^2}).$$

考虑到控制集 U 的结构, 不难计算, 极大化控制 $\widehat{u}_N(\cdot)$ 依赖参数 $\lambda \in (0,1)$ 如下:

(a) 如果 $\lambda > 1/\sqrt{17}$, 那么 $\widehat{u}_{1N}(t) = 2$, $\widehat{u}_{2N}(t) = 2$ 对所有 $t \in T_N$ 成立;

(b) 如果 $\lambda < 1/\sqrt{17}$, 那么 $\widehat{u}_{1N}(t) = 0$, $\widehat{u}_{2N}(t) \in [0,2]$ 对所有 $t \in T_N$ 成立;

(c) 如果 $\lambda = 1/\sqrt{17}$, 那么对每个 $t \in T_N$, 或者 $\widehat{u}_{1N}(t) = \widehat{u}_{2N}(t) = 2$, 或者 $\widehat{u}_{1N}(t) = 0$ 和 $\widehat{u}_{2N}(t) \in [0,2]$.

可直接验证, 情形 (a) 中的控制 $\widehat{u}_N(\cdot)$ 对 (P_N) 不是可行的, 这是因为对应的轨道 $\widehat{x}_N(\cdot)$ 不满足端点约束. 在情形 (b) 中, 控制 $\widehat{u}_N(\cdot)$ 与最优性相去甚远, 这是因为 $J_N[\widehat{u}_N, \widehat{x}_N] = 0$, 但 $\inf J[u_N, x_N] \leqslant -1$. 在情形 (c) 中, 当 $N \in \mathbb{N}$ 时, 如果使 $\widehat{u}_{1N}(t) = 2$ 且 $\widehat{u}_{2N}(t) = 2$ 的点 $t \in T_N$ 的数量不大于 $[N/2]$, 那么控制 $\widehat{u}_N(\cdot)$ 对 (P_N) 是可行的. 由定理 6.59 和现在这个例子之前的讨论可断言, (P_N) 的最优控制 $\bar{u}_N(\cdot)$ (总是存在的) 可能是情形 (c) 中满足恰当条件的可行控制 $\widehat{u}_N(\cdot)$, 或者是那些使费用和约束函数的值与 $\varphi_0(\widehat{x}_N(b))$ 和 $\varphi_1(\widehat{x}_N(b))$ 相差 $o(h_N)$ 数量级 (当 $N \to \infty$ 时) 的控制.

因此, AMP 可用来有效地描述 (P_N) 具有最优性可能的所有可行控制集合. 根据这一信息, 最后由问题 (P_N) 的结构可确定, 这些问题序列的最优解由如下的控制给出

$$\begin{cases} \bar{u}_{1N}(t) = \bar{u}_{2N} = 2, & \text{如果 } t \text{ 是 } T_N \text{ 中的第 } [\gamma_N/2h_N] \text{ 个点}, \\ \bar{u}_{1N}(t) = 0, \quad \bar{u}_{2N}(t) \in [0,2], & \text{对所有其他 } t \in T_N. \end{cases}$$

这就完全解决了所考虑的问题.

6.4.6 时滞和中立型控制系统

本小节致力于研究 AMP 在具有光滑动态时滞控制系统有限差分逼近的上次微分形式中的推广. 为了简洁起见, 只给出自由端点问题的结果. 本小节的主要定理是定理 6.50 在时滞问题情形的一个推广, 定理 6.59 和定理 6.66 对应的推广可类似导出. 另一方面, 在本小节最后证明了, 对于不仅关于状态变量而且关于速率变量都包含时滞的中立型光滑泛函–微分系统的离散逼近, AMP 可能不成立.

从下面的在可测控制 $u\colon [a,b] \to U$ 和对应时滞系统的绝对连续轨道 $x\colon [a,b] \to X$ 上, 关于状态变量具有单个时滞的连续时间问题 (D) 开始:

$$\begin{cases} \min & J[u,x] := \varphi(x(b)) \\ \text{s.t.} & \dot{x}(t) = f(x(t), x(t-\theta), u(t), t), \quad \text{a.e. } t \in [a,b], \\ & x(t) = c(t), \quad t \in [a-\theta, a], \\ & u(t) \in U, \quad \text{a.e. } t \in [a,b], \end{cases}$$

其中 $\theta > 0$ 为常数时滞且 $c\colon [a-\theta, a] \to X$ 为给定的函数, 它用来定义启动时滞系统所需要的初始“尾”条件; 参见注 6.40, 在那里, 对这类问题关于最大值原理的结果已经讨论过. 现在的目标是, 对时滞问题 (D) 的离散逼近, 导出适当形式的 AMP.

由 Euler 有限差分替换导数, 下面构造 (D) 的离散逼近. 在时滞问题的情形, 需要确保, 只要 t 属于离散格点, 那么点 $t-\theta$ 也属于离散格点. 这可通过定义离散步长 $h_N := \dfrac{\theta}{N}$ 来实现, 这不同于 6.4.3 小节中对非时滞问题 (P_N^0) 使用的 $h_N = \dfrac{b-a}{N}$. 在这个方案中, 时间区间长度 $b-a$ 与离散化步长 h_N 一般不可公度. 在主要时间区间 $[a,b]$ 上定义格点 T_N 为

$$T_N := \{a, a+h_N, \cdots, b-\widetilde{h}_N - h_N\}, \quad \text{其中}$$
$$h_N := \frac{\theta}{N}, \quad \widetilde{h}_N := b - a - h_N\left[\frac{b-a}{h_N}\right],$$

并且考虑下面具有离散时滞的有限差分逼近问题序列 (D_N):

$$\begin{cases} \min & J[u_N, x_N] := \varphi(x_N(b)) \\ \text{s.t.} & x_N(t+h_N) = x_N(t) + h_N f(x_N(t), x_N(t-Nh_N), u_N(t), t), \ t \in T_N, \\ & x_N(b) = x_N(b-\widetilde{h}_N) + \widetilde{h}_N f(x_N(b-\widetilde{h}_N), u_N(b-\widetilde{h}_N), b-\widetilde{h}_N), \\ & x_N(t) = c(t), \quad t \in T_{0N} := \{a-\theta, a-\theta+h_N, \cdots, a\}, \\ & u_N(t) \in U, \quad t \in T_N. \end{cases}$$

为了导出问题序列 (D_N) 的 AMP, 将这些问题约化为无时滞问题, 利用定理 6.57 的结果, 其中的常驻假设类似 6.4.3 小节中阐述的假设, 现在在 $f(x,y,u,t)$ 中

涉及除 x 外的状态变量 y. 为了方便起见, 引入下面的记号:

$$z_N(t) := (x_N(t), x_N(t-\theta)), \quad \bar{z}_N(t) := (\bar{x}_N(t), \bar{x}_N(t-\theta)),$$
$$f(z_N, u_N, t) := f(x_N(t), x_N(t-\theta), u_N(t), t),$$
$$f(\bar{z}_N, u_N, t) := f(\bar{x}_N(t), \bar{x}_N(t-\theta), u_N(t), t).$$

那么对每个 $N \in \mathbb{N}$, 沿着时滞问题 (D_N) 的最优过程 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$, (D_N) 的伴随系统可写为

$$\begin{aligned} p_N(t) =& p_N(t+h_N) + h_N \nabla_x f(\bar{z}_N, \bar{u}_N, t)^* p_N(t+h_N) \\ &+ h_N \nabla_y f(\bar{z}_N, \bar{u}_N, t+\theta)^* p_N(t+\theta+h_N) \quad (t \in T_N), \\ p_N(b-\widetilde{h}_N) =& p_N(b) + \widetilde{h}_N \nabla_x f(\bar{z}_N, \bar{u}_N, b-\widetilde{h}_N)^* p_N(b). \end{aligned}$$

引入对应的 Hamilton-Pontryagin 函数

$$H(x_N, y_N, p_N, u, t) := \begin{cases} \langle p_N(t+h_N), f(x_N, y_N, u, t)\rangle, & \text{如果 } t \in T_N, \\ \langle p_N(t), f(x_N, y_N, u, t-\widetilde{h}_N)\rangle, & \text{如果 } t = b-\widetilde{h}_N, \end{cases}$$

其中 $y_N(t) := x_N(t-\theta)$. 那么伴随系统在 $t \in T_N$ 时可写为

$$p_N(t) = p_N(t+h_N) + h_N\Big[\nabla_x H(\bar{z}_N, p_N, \bar{u}_N, t) + \nabla_y H(\bar{z}_N, p_N, \bar{u}_N, t+\theta)\Big],$$

在“不可公度”点可写为

$$p_N(b-\widetilde{h}_N) = p_N(b) + \widetilde{h}_N \nabla_x H(\bar{z}_N, p_N, \bar{u}_N, b-\widetilde{h}_N).$$

于是有下面关于满足时滞离散逼近 AMP 的结果.

定理 6.69(时滞系统的 AMP) 设 $\{\bar{u}_N(\cdot), \bar{x}_N(\cdot)\}$ 为问题 (D_N) 的最优解. 除了所作的常驻假设外, 假设费用函数 φ 在序列 $\{\bar{x}_N(b)\}$ $(N \in \mathbb{N})$ 的极限点附近是一致上次可微的. 那么对每个上次梯度序列 $x_N^* \in \mathcal{D}^+\varphi(\bar{x}_N(b))$, 近似最大值条件

$$H(\bar{z}_N, p_N, \bar{u}_N, t) = \max_{u \in U} H(\bar{z}_N, p_N, u, t) + \varepsilon(t, h_N), \quad t \in \widetilde{T}_N := T_N \cup \{b-\widetilde{h}_N\}$$

得到满足, 其中当 $h_N \to 0$ 时, $\varepsilon(t, h_N) \to 0$ 关于 $t \in \widetilde{T}_N$ 一致, 并且 $p_N(\cdot)$ 满足横截性条件

$$p_N(b) = -x_N^*, \quad p_N(t) = 0 \quad (\text{当 } t > b \text{ 时}). \tag{6.105}$$

更进一步, 如果 X 是自反的且 φ 在 $\{\bar{x}_N(b)\}$ 的极限点附近是连续的, 那么在 (6.105) 中可选取任何 $x^* \in \widehat{\partial}^+\varphi(\bar{x}_N(b))$.

证明 下面分成几步将时滞离散逼近问题约化为无时滞问题 (但 $b-a$ 和 h_N 不可公度). 记

$$
\begin{aligned}
&y_{1N}(t) := x_N(t-h_N), \quad t \in \{a+2h_N, \cdots, b-\widetilde{h}_N\},\\
&y_{1N}(t) := c_N(t-h_N), \quad t \in \{a-\theta+h_N, \cdots, a+h_N\},\\
&y_{2N}(t) := y_{1N}(t-h_N), \quad t \in \{a-\theta+2h_N, \cdots, b-\widetilde{h}_N\},\\
&\quad\vdots\\
&y_{NN}(t) := y_{N-1,N}(t-h_N), \quad t \in \{a, \cdots, b-\widetilde{h}_N\},
\end{aligned}
$$

并且观察到 $y_{1N}(b), \cdots, y_{NN}(b)$ 的值可任意定义, 因为它们既不出现在伴随系统中, 也不出现在费用函数中. 为了与定理 6.57 的设置相配, 定义

$$
y_{1N}(b) := x_N(b-\widetilde{h}_N),\ y_{2N}(b) := y_{1N}(b-\widetilde{h}_N), \cdots, y_{NN}(b) := y_{N-1,N}(b-\widetilde{h}_N).
$$

经过变量变换后有

$$
y_{NN}(t) = \begin{cases} x_N(t-\theta), & t \in \{a+\theta+h_N, \cdots, b-\widetilde{h}_N\},\\ c(t-\theta), & t \in \{a, \cdots, a+\theta\}. \end{cases}
$$

(D_N) 中的原始系统, 对每个 $N \in \mathbb{N}$, 由此可约化为下面 $\mathbb{R}^{(N+1)n}$ 维无时滞的系统:

$$
\begin{cases} s_N(t+h_N) = s_N(t) + h_N g(s_N, u_N, t), \quad t \in T_N,\\ s_N(b) = s_N(b-\widetilde{h}_N) + \widetilde{h}_N g(s_N, u_N, b-\widetilde{h}_N), \end{cases}
$$

其中, 状态向量 $s_N(t) := (x_N(t), y_{1N}(t), \cdots, y_{NN}(t))$, 并且“速率”映射 $g(s_N, u_N, t)$ 为

$$
g(s_N(t), u_N(t), t) = \begin{cases} f(x_N(t), y_{NN}(t), u_N(t), t),\\ \dfrac{x_N(t) - y_{1N}(t)}{h_N},\\ \quad\vdots\\ \dfrac{y_{N-1,N}(t) - y_{NN}(t)}{h_N}, \end{cases}
$$

在最后这个公式中, 当 $t = b - \widetilde{h}_N$ 时, h_N 应当替换为 $\widetilde{h}_N$.

在所得到的无时滞系统的可行对 $\{u_N(\cdot), s_N(\cdot)\}$ 上, 将定理 6.57 应用到极小化和 (D_N) 中相同的泛函问题上. 该问题关于新的伴随变量 $q \in \mathbb{R}^{(N+1)n}$ 的伴随系统具有形式

$$
\begin{cases} q_N(t) = q_N(t+h_N) + h_N \nabla_s g(\bar{s}_N, \bar{u}_N, t)^* q(t+h_N), \quad t \in T_N,\\ q_N(b-\widetilde{h}_N) = q_N(b) + \widetilde{h}_N \nabla_s g(\bar{s}_N, \bar{u}_N, b-\widetilde{h}_N)^* q_N(b), \end{cases}
$$

以及横截性条件

$$q_N(b) = -(x_N^*, 0, \cdots, 0)$$

对 $x_N^* \in \mathcal{D}^+\varphi(\bar{x}_N(b))$ 成立. 当 X 是自反且 φ 连续时, 它可约化为 $x_N^* \in \widehat{\partial}^+\varphi(\bar{x}_N(b))$. 考虑到上面 g 和 f 的关系并且由初等计算, 可将算子 $\nabla_s g^*$ 通过 $\nabla_x f^*$ 和 $\nabla_y f^*$ 表示, 并且对伴随轨道 $q_N(\cdot)$ 的第一个分量 $p_N(\cdot)$ 得到横截性条件 (6.105). 更进一步, 可得到上述非时滞和原时滞系统 Hamilton-Pontryagin 函数的关系式

$$\begin{aligned}\widetilde{H}(\bar{s}_N, q_N, u, t) &= \langle q_N(t+h_N), g(\bar{s}_N, u, t)\rangle \\ &= \langle p_N(t+h_N), f(\bar{z}_N, u, t)\rangle + r(\bar{s}_N, q_N, h_N, t) \\ &= H(\bar{z}_N, p_N, u, t) + r(\bar{s}_N, q_N, h_N, t), \quad t \in T_N,\end{aligned}$$

若 $t = b - \widetilde{h}_N$, 则有类似关系, 其中余项 $r(\bar{s}_N, q_N, h_N, t)$ 不依赖 u. 现在将定理 6.57 的近似最大值原理应用到无时滞系统上, 完成了定理的证明. △

作为本节的结束, 考虑所谓的中立型泛函微分系统 (也请比照 7.1 节) 的有限差分近似最优控制问题, 该中立型泛函微分系统由

$$\dot{x}(t) = f(x(t), x(t-\theta), \dot{x}(t-\theta), u(t), t), \quad u(t) \in U, \quad \text{a.e. } t \in [a, b]$$

给出, 它不仅在状态变量中, 而且也在速率变量中包含时滞. 这类系统的具有步长 h 和格点 $T := \{a, a+h, \cdots, b-h\}$ 的有限差分版本为

$$x(t+h) = x(t) + hf\left(x(t), x(t-\theta), \frac{x(t-\theta+h) - x(t-\theta)}{h}, u(t), t\right),$$

其中 $u(t) \in U$ $(t \in T)$ 并且伴随系统为

$$\begin{aligned}p(t) = {} & p(t+h) + h\nabla_x f(\bar{v}, \bar{u}, t)^* p(t+h) + h\nabla_y f(\bar{v}, \bar{u}, t+\theta)^* p(t+\theta+h) \\ & + h\nabla_z f(\bar{v}, \bar{u}, t+\theta-h)^* p(t+\theta) - h\nabla_z f(\bar{v}, \bar{u}, t+\theta)^* p(t+\theta+h)\end{aligned}$$

对 $t \in T$ 成立, 其中 $\{\bar{u}(\cdot), \bar{x}(\cdot)\}$ 为问题 (D_N) 的中立型变体的最优解, 并且

$$\bar{v}(t) := \left(\bar{x}(t), \bar{x}(t-\theta), \frac{\bar{x}(t-\theta+h) - \bar{x}(t-\theta)}{h}\right), \quad t \in T.$$

下面的例子说明, AMP 对有限差分中立型系统一般是不成立的, 即使具有光滑费用函数的时候亦如此, 这与普通或时滞系统是不同的.

例 6.70(AMP 对中立型系统可能不成立) 在无端点约束的光滑中立系统上, 存在极小化线性函数的二维控制问题, 使得无论步长和网格点如何选取, 离散逼近的某最优控制序列不满足近似最大值原理.

证明 考虑下面具有参数 $h>0$ 的中立型系统离散最优控制问题的参数族:

$$
\begin{cases}
\min & J[u,x_1,x_2]:=x_2(2)\\
\text{s.t.} & x_1(t+h)=x_1(t)+hu(t),\quad t\in T:=\{0,h,\cdots,2-h\},\\
& x_2(t+h)=x_2(t)+h\left(\dfrac{x_1(t-1+h)-x_1(t-1)}{h}\right)^2-hu^2(t),\quad t\in T,\\
& x_1(t)\equiv x_2(t)\equiv 0,\quad t\in T_0:=\{-1,\cdots,0\},\\
& |u(t)|\leqslant 1,\quad t\in T.
\end{cases}
$$

易见

$$x_2(1)=-h\sum_{t=0}^{1-h}u^2(t),$$

$$
\begin{aligned}
x_2(2)&=x_2(1)+h\sum_{t=1}^{2-h}\left(\frac{x_1(t-1+h)-x_1(t-1)}{h}\right)^2-h\sum_{t=1}^{2-h}u^2(t)\\
&=-h\sum_{t=0}^{1-h}u^2(t)+h\sum_{t=0}^{1-h}u^2(t)-h\sum_{t=1}^{2-h}u^2(t)=-h\sum_{t=1}^{2-h}u^2(t).
\end{aligned}
$$

因此, 对任何 h, 控制

$$\bar{u}(t)=\begin{cases}0, & t\in\{0,\cdots,1-h\},\\ 1, & t\in\{1,\cdots,2-h\}\end{cases}$$

是所考虑问题的一个最优控制. 对应的轨道为

$$\bar{x}_1(t)=\begin{cases}0, & t\in\{0,\cdots,1-h\},\\ t-1, & t\in\{1,\cdots,2-h\};\end{cases}\qquad \bar{x}_2(t)=\begin{cases}0, & t\in\{0,\cdots,1-h\},\\ -t+1, & t\in\{1,\cdots,2-h\}.\end{cases}$$

在上面的系统中, 计算"速率"映射 f 的偏导数, 则有

$$\nabla_x f=\begin{pmatrix}0&0\\0&0\end{pmatrix},\quad \nabla_y f=\begin{pmatrix}0&0\\0&0\end{pmatrix},$$

$$\nabla_z f(t+1)=\frac{1}{h}\begin{pmatrix}0&0\\2(x_1(t+h)-x_1(t))&0\end{pmatrix}.$$

因此, 伴随系统约化为

$$
\begin{aligned}
p_1(t)=&\,p_1(t+h)+2(\bar{x}_1(t)-\bar{x}_1(t-h))p_2(t+1)\\
&-2(\bar{x}_1(t+h)-\bar{x}_1(t))p_2(t+1+h),\quad t\in\{0,\cdots,2-h\},
\end{aligned}
$$

$p_2(t)$ 恒为常数并且具有横截性条件

$$p_1(2) = 0, \quad p_2(2) = -1; \quad p_1(t) = p_2(t) = 0 \quad (t > 2).$$

该系统的解为

$$p_1(t) \equiv 0, \quad p_2(t) \equiv -1 \quad \text{对所有} \quad t \in \{0, \cdots, 2-h\}.$$

因此, Hamilton-Pontryagin 函数沿着最优解为

$$\begin{aligned} H(t, \bar{x}_1, \bar{x}_2, p_1, p_2, u) = p_2(t+h) \left\{ \left(\frac{x_1(t-1+h) - x_1(t-1)}{h} \right)^2 - u^2 \right\} \\ + p_1(t+h)u = u^2, \quad t \in \{0, \cdots, 1-h\}. \end{aligned}$$

这证明了, 不管 h 和网格点 $t \in \{0, \cdots, 1-h\}$ 如何选取, 最优控制 $\bar{u}(t) = 0$ 没有给 Hamilton-Pontryagin 函数提供近似极大值. 同时注意到, 另一个最优控制序列 $\bar{u}(t) = 1$ $(\forall t \in \{0, \cdots, 2-h\})$, 无论 h 如何选取都不满足确切离散最大值原理. △

6.5 第 6 章的评注

6.5.1 变分法与最优控制

第 6 章致力于动态最优化问题的研究. 动态最优化这个名称通常反映了最优化问题的某些初始数据随着时间的改变而变化这一事实. 这类问题的起源可追溯到经典的变分法, 而变分法则是所有无穷维分析的起点; 在这方面的发展中, 读者可参阅对本书涉及课题影响最大的开创性的贡献, 包括 Euler [411], Lagrange [737], Hamilton [548], Jacobi [625], Mayer [859], Weierstrass [1326], Bolza [130], Tonelli [1260], Carathéodory[222] 和 Bliss [119] (以及他的著名的 Chicago 学派的著作) 等.

常微分方程 (ODE) 的最优控制理论作为经典的变分法的一个现代版本为人们所熟知. 它与变分法的首要区别在于控制函数上出现的硬/逐点约束, 由之通过发展型 ODE 系统

$$\dot{x} = f(x, u, t), \quad u(t) \in U, \quad t \in [a, b], \quad x \in \mathbb{R}^n, \tag{6.106}$$

生成了系统的轨道 (通常称之为容许轨线). 这类由集合 U 给出的控制约束具有相当不规则特性, 这在产生于实际问题中的最早的最优控制问题中就出现了, 一直就是最优控制问题内蕴非光滑性的固有根源, 并最终推动了现代变分分析和广义微分诸多关键方面发展.

如 1.4.1 小节中所述, 人们熟知的 Pontryagin 最大值原理 (PMP)[1102] 是最优控制理论的基本结果, 在现代变分分析的发展中起到了重要的作用. 它最初是由

Pontryagin 表述的并由 Gamkrelidze[494] 对线性的情形和 Boltyanskii[124] 对具有非线性光滑动态的情形进行证明. 有趣的是, 作为局部最优性的一个充分条件, 最大值原理表述的初次尝试 [129] 是错误的; 关于最大值原理发现, 参见 Boltyanskii[128] 和 Gamkrelidge[498] 文章中的史料 (二者相当不同). 在这些文章中, Hestenes[565] 的著作以及 McShane[865] 的综述文章中, 读者可以找到关于最大值原理和以前 Chicago 学派关于变分法以及在自动控制的理论和应用中所得结果之间关系的各种讨论; 也请参见 Gabasov 与 Kirillova 的绝妙综述 [487]. 在最优控制理论之前, 也许离它最近的要数非标准变分问题以及线性自动控制最优系统中的结果, 尤其是, 由 Feldbaum[440] 得到的所谓“n- 区间定理”以及 Bellman, Glicksberg 与 Gross[95] 得到的“砰–砰原理”.

虽然 PMP 的阐述和证明中的许多要素在变分法中可以找到类似的对应 (这特别包括, McShane[860] 使用的针形变分, 这实际上得追溯到 Weierstrass[1326] 和他的强极小点的必要最优性条件; 切向凸逼近和如在 McShane[860] 中的凸分离的应用; 典则变量和修改的 Hamilton 函数等), 但是 PMP 的发现及其证明仍是令人惊喜的 (用 Pshenichnyi[1106] 的话来说, 是“轰动”的). 在现代变分分析的发展过程中, 无法估量 PMP 的影响和地位. 关于最优控制, 变分法和数学规划之间的关系的更多结果和讨论, 请读者参阅文献 [7, 32, 105, 124, 218, 235, 255, 370, 485,486, 497, 504, 539, 565, 618, 801, 863, 865, 877, 1002, 1106, 1239, 1289, 1315, 1351].

与经典的变分法相比, (6.106) 的伴随系统

$$\dot{p} = -\frac{\partial f(\bar{x}, \bar{u}, t)^*}{\partial x} p = -\nabla_x H(\bar{x}, p, \bar{u}, t) \tag{6.107}$$

的发现 (归功于 Pontryagin) 似乎是 PMP 最为显著的新贡献之一, 它是通过 Hamilton-Pontryagin 函数

$$H(x, p, u, t) := \langle p, f(x, u, t) \rangle, \quad p \in \mathbb{R}^n, \tag{6.108}$$

沿着最优过程 $(\bar{x}, \bar{u})$ 计算的; 由此关键的逐点最大值条件可写为

$$H\big(\bar{x}(t), p(t), \bar{u}(t), t\big) = \max_{u \in U} H\big(\bar{x}(t), p(t), u, t\big), \quad \text{a.e..} \tag{6.109}$$

在 PMP 被发现之后, 人们发现最大值条件 (5.109) 可看成是变分法中强极小点的 Weierstrass 的剩余函数条件相对应的最优控制版本.

6.5.2 微分包含

原始最优控制模型 (1.106) 的一个显著不足之处在于它无法涵盖具有依赖于状态的控制集 $U = U(x)$ 的问题, 该类问题对理论和应用都是重要的. 这类问题以及

控制和动态最优化中其他重要的问题能够自然地写为微分包含的形式

$$\dot{x} \in F(x,t), \quad x \in \mathbb{R}^n. \tag{6.110}$$

实际上, 这可追溯到 20 世纪 30 年代对集值微分方程的研究 (不是从控制的角度), 如 Marchaud [850] 的"相依切方程"和 Zaremba [1355] 的"拟切方程"; 其早期的发展也可参见 Nagumo [990] 和 Wazewsk [1325]. 由所谓的"Filippov 隐函数引理"[449] (它事实上是关于集值映射可测选择的一个结果), 控制系统 (6.106) 可等价地约化为微分包含 (6.110) 的形式; 更多的参考文献和讨论可参见例如 Castaing 与 Valadier [229] 以及 Rockafellar 与 Wets [1165].

注意到, 与 (6.106) 的经典情形相比, 由微分包含 (6.110) 控制的控制系统要复杂得多. 这是由于, 例如, 在此情形中不可能利用标准的针形变分来导出最优性条件. 此外, 系统 (6.110) 显式地揭示了内蕴的非光滑性, 即使在经典的最优控制中也是固有的, 这首先是由 $u(t) \in U$ 类型的硬控制约束造成的, 特别是 $U = \{0,1\}$ 这样的有限集的情形, 这在自动控制的应用中是比较典型的. 由于使用了关于状态 - 余状态变量 (x,p) 可微的 Hamilton-Pontryagin 函数 (6.108), 该现象在光滑动态系统 (6.106) 的 PMP 中是隐含的. 最优控制中非光滑性还显示在微分包含 (6.110) 的 Hamilton 函数上, 即

$$\mathcal{H}(x,p,t) := \sup\{\langle p,v\rangle \ \mid u \in F(x,t)\}. \tag{6.111}$$

该 Hamilton 函数对应于标准/参数化控制系统 (6.106) 的原 (真正) Hamilton 函数

$$\mathcal{H}(x,p,t) := \sup\{H(x,p,u,t) \mid v \in U\}.$$

这些广义 Hamilton 函数可视为变分法和力学问题中经典 Hamilton 函数对应的控制版本. 而经典的 Hamilton 函数涉及 Lagrange 函数, 即在极小化下的被积函数 (在良好定义的时候, 它可由 Lagrange 函数通过 Legendre 变换得到).

6.5.3　光滑或图凸 (graph-convex) 微分包含的最优性条件

非光滑性是 Hamilton 函数 (6.111) 及其在上述控制系统 (6.106) 上的具体实现的一个标志性特征; 光滑的情形只有在某些相当严格的假设下才出现. 然而, 由微分包含控制的控制问题的第一个必要最优性条件是由 Boltyanskii[125] 在 (6.111) 的关于状态变量光滑性的假设下得到的 (冠以"支撑原理"的名称); 也参见相关的文章 Fedorenko [438, 439], Boltyanskii [127], Blagodatskikh [117], Blagodatskikh 与 Filippov [118] 以及其参考文献 (大多是俄语).

在文献 [1143–1145] 中, Rockafellar 使用了更加合理的假设, 即 $F(\cdot,t)$ 的图凸性 (graph-convexity), 导出了微分包含 (6.110) 的必要 (和充分) 最优性条件. 实际

上, Rockafellar 考虑了更加广泛的 (完全) 凸的广义 Bolza 问题:

$$\min \quad \varphi\big(x(a), x(b)\big) + \int_a^b \vartheta\big(x(t), \dot{x}(t), t\big)\mathrm{d}t, \tag{6.112}$$

其中对比经典 Bolza 问题 [130] 和先前 $\vartheta = 0$ 情形的 Mayer 问题 [859], 函数 φ 和 ϑ 可能是增广实值的, 即 (6.112) 特别地将微分包含 (6.110) 通过指标函数 $\vartheta(x, v, t) := \delta((x, v); \mathrm{gph}\, F(\cdot, t))$ 纳入其中. 在文献 [1143–1145] 中, 所作的 $\vartheta(x, v, t)$ 关于变量 (x, v) 的凸性假设蕴涵着涉及微分包含 (6.110) 的 Hamilton 函数 (6.111) 关于 p 是凸的并且关于 x 是凹的, 因而, 在凸分析的意义下, 作为鞍函数, 它关于 (x, p) 是次可微的. 由此 Rockafellar 利用无穷维空间中凸分析的机理得到了凸广义 Bolza 问题最优解 $\bar{x}(\cdot)$ 的必要和充分条件, 并因此对凸图 (convex-graph) 微分包含的情形通过广义 Hamilton 方程 [1145](也称为 Hamilton 条件/包含)

$$\big(-\dot{p}(t), \dot{\bar{x}}(t)\big) \in \partial\mathcal{H}\big(\bar{x}(t), p(t), t\big), \quad \text{a.e.} \tag{6.113}$$

得到最优解 $\bar{x}(\cdot)$ 的必要和充分条件, 其中 $\partial\mathcal{H}$ 表示 Hamilton 函数 $\mathcal{H}(x, p, t)$ 关于 (x, p) 的次微分. 如果 $\mathcal{H}(x, p, t)$ 关于 x 和 p 是可微的, 那么包含 (6.113) 可约化为经典的 Hamilton 系统

$$\dot{\bar{x}}(t) = \nabla_p\mathcal{H}\big(\bar{x}(t), p(t), t\big), \quad -\dot{p}(t) = \nabla_x\mathcal{H}\big(\bar{x}(t), p(t), t\big).$$

由凸图微分包含控制的最优化问题的有点不同的结果 (但大多是等价的) 后来由 Halkin[542], Berliocchi 与 Lasry[107], 以及 Pshenichnyi[1107, 1109] 得到.

6.5.4 Clarke 的 Euler-Lagrange 条件

注意到, 在与 Hamilton 函数上的光滑性要求相比较, 虽然在 $F(\cdot, t)$ 上的图凸性假设更合理, 但仍然相当苛刻. 尤其, 对标准控制系统 (6.106), 该假设实际上可约化为 $f(\cdot, \cdot, t)$ 的线性性和 U 的凸性; 参见 Rockafellar [1143]. 从完全凸 (或采用 Halkin 的术语“双凸”) 问题 (即 (6.112) 中被积函数关于 (x, v) 是凸的) 到只涉及关于速率变量 v 的凸性 (这在微分包含的框架中对应于 $F(x, t)$ 的凸值性) 问题, 关键的一步是 Clarke 于 20 世纪 70 年代从他的博士论文 [243] 开始的开创性工作.

Clarke[243, 245] 的出发点是 Bolza 型问题 (6.112), 它具有有限 (事实上是 Lipschitz) 的被积函数/Lagrange 函数 $\vartheta(\cdot, \cdot, t)$, 在被积函数 ϑ 和可取无限值的增广实值的下半连续 (l.s.c) 端点函数 φ 上无任何光滑性和凸性假设. 主要的必要最优性条件是通过 $\vartheta(\cdot, \cdot, t)$ 的 Clarke 广义梯度以 Euler-Lagrange 的形式

$$(\dot{p}(t), p(t)) \in \partial_C\vartheta(\bar{x}(t), \dot{\bar{x}}(t), t), \quad \text{a.e.} \tag{6.114}$$

得到的. 如果 $\vartheta(x,v,t)$ 关于 (x,v) 是光滑的, 那么包含 (6.114) 回到经典的 Euler-Lagrange 方程; 如果 ϑ 关于 x 和 v 都是凸的, 那么它约化为 Rockafellar 于文献 [1143] 得到的 Euler-Lagrange 包含. 更进一步可以看到, 在文献 [243, 245] 中, Clarke 对 (6.114) 的证明是基于将非凸 Bolza 问题约化为 Rockafellar 全面研究过的完全凸问题. 在该约化过程进而在 (6.114) 的整个证明中, Clarke 广义梯度的凸值性及其与它的广义方向导数的对偶关系的作用是决定性的.

基于有限 Lagrange 函数的 Euler-Lagrange 条件 (6.114), Clarke 在文献 [247] 中通过 $F=F(\cdot,t)$ 图像的 Clarke 法锥对 Lipschitz 和有界微分包含 (6.110) 得到了相应的包含

$$(\dot{p}(t),p(t))\in N_C\big((\bar{x}(t),\dot{\bar{x}}(t));\operatorname{gph}F(t)\big),\quad \text{a.e.}. \tag{6.115}$$

然后, 他在文献 [248] 中导出了广义 Bolza 问题 (6.112) 的 Euler-Lagrange 包含 (6.114), 其中 ϑ 是增广实值的并且关于 (x,v) 是上图 Lipschitz 的. 在文献 [247, 248] 中, 最为显著和苛刻的假设是平静性条件, 它类似 5.5.16 小节中讨论过的数学规划问题中的条件. 这是一种约束规范/正则性要求, 它确保了标准/规范形式的必要最优性条件, 尤其当端点函数 φ 关于任何一个变量是局部 Lipschitz 时它是成立的; 然而, 此时它排除了对应的端点约束. 注意到, 由于平静性的要求, Clarke 在形式上避免了在 ϑ (甚至于关于 v) 上的凸性假设, 但由于平静性提供的 "容许松弛", 凸性性质实际上还是在文献 [247,248] 中出现了; 有关这些关系的详细研究亦可参见文献 [246]. 进一步须指出, 如文献 [248, 第 683 页] 中提到的, "…… 由 Rockafellar 发展的双凸情形是这些结果的证明的核心."

(6.115) 形式的 Euler-Lagrange 包含的最为严重的缺陷在于它涉及 (6.110) 中 $F(\cdot,t)$ 图像的 Clarke 法锥, 只要 F 在最优解附近是图像 Lipschitz 的, 则该法锥是维数 $d\geqslant n$ 的一个线性子空间, 这一缺陷是在后来才充分认识到的; 更多的讨论请参见 1.4.4 小节. 由于这个性质, (6.115) 右端的集合可能太大, 在理论和应用的许多重要情形中无法给伴随轨线 $p(\cdot)$ 提供足够的信息.

6.5.5 Clarke 的 Hamilton 条件

除了 Euler-Lagrange 条件 (6.115) 外, Clarke 也建立了广义 Bolza 问题的必要最优性条件, 并因此建立了 Lipschitz 微分包含的如下 Hamilton 形式

$$(-\dot{p}(t),\dot{\bar{x}}(t))\in\partial_C\mathcal{H}(\bar{x}(t),p(t),t),\quad \text{a.e.}. \tag{6.116}$$

它涉及了 Hamilton 函数关于变量 (x,p) 的 Clarke 广义梯度. 其中第一个 Hamilton 结果是在平静性假设下得到的 [253, 255], 然后在无该假设和其他约束规范条件下得到这一结果 [256].

注意到, 在没有正则性/正规性的假设下, Hamilton 条件 (6.116) 只是对凸值微分包含建立的 (它对应于广义 Bolza 形式中 Lagrange 函数关于 v 是凸的情形); 最初, 文献 [251] 第 265 页中给出的关于 (6.116) 无凸性论断的证明是不正确的, 该证明关系到凸化过程. 然而, 利用类似的过程, 基于 Ekeland 变分原理来证明非光滑最优控制系统 (6.106) 的 Pontryagin 最大值原理的 Clarke 推广 [250] 是可行的. 在没有上述凸性的假设下, 关于 Hamilton 必要最优性条件 (6.116) 的正确性曾是存在很长时间的一个猜想, 不少人尝试过都没有成功; 最近 Clarke 在 [261] 中解决了 Lipschitz 和有界微分包含的情形, 其中利用了 Stegall 变分原理而不是 Ekeland 变分原理. 注意到, 与经典光滑情形和 Rockafellar 的完全凸情形不同, Clarke 的 Euler-Lagrange 条件 (6.115) 和 Hamilton 条件 (6.116), 既使是在简单的情形下, 也是不等价的. 此外, 从其中之一得不到另外一个, 因为它们真是相互独立的; 参见 Kaskosz 与 Lojasiewicz [667] 以及 Loewen 与 Rockafellar [805] 中的例子和讨论.

直到 Loewen 与 Rockafellar[804] 的工作, 人们才清楚是否能找到一个共同的伴随轨线 $p(\cdot)$ 同时满足 Euler-Lagrange 条件 (6.115) 和 Hamilton 条件 (6.116). 在没有平静性或正规性的假设下, 对于凸值和 Lipschitz 微分包含, 文献 [804] 给出了肯定的答案. 注意到, 在该情形中, 条件 (6.115) 和 (6.116) 自动地蕴涵着 Weierstrass-Pontryagin 最大值条件

$$\langle p(t), \dot{\bar{x}}(t)\rangle = \mathcal{H}(\bar{x}(t), p(t), t), \quad \text{a.e.} \tag{6.117}$$

对于动态最优化和最优控制的各种问题, 由 Clarke 广义微分结构描述的 Euler-Lagrange 和 Hamilton 类型的必要最优性条件的推广和修改, 推荐读者参阅文献 [254–256, 267, 268, 272–274, 276, 595, 666, 667, 803, 804, 808, 1178, 1291, 1292] 以及其中的文献.

6.5.6　横截性条件

除了上面讨论的动态关系 (Euler-Lagrange, Hamilton, Weierstrass-Pontryagin 等条件), 动态最优化问题的必要最优性条件也包括称为横截性条件的伴随轨道上的端点关系. 它们是通过依赖于状态轨道端点的费用和约束函数的适当 (广义) 微分结构表示的. 注意到, 如果端点费用函数 φ 像广义 Bolza 问题 (6.112) 中那样假设为增广实值的话, 那么 $(x(a), x(b))$ 上的端点约束可隐含地包括在 φ 中. 然而, 通常这类约束由

$$(x(a), x(b)) \in \Omega \subset \mathbb{R}^n \tag{6.118}$$

的形式显式地给出, 其中约束/目标集合 Ω 可以以某些泛函的形式具体给出, 例如实值 (通常是 Lipschitz) 函数的等式和不等式形式.

上面提到的 Clarke 和他的追随者发表的文章中, 所考虑的是 (6.112) 中的极小化 Lipschitz 费用函数 φ 受限于 (6.118) 类型的端点约束问题. 在最优端点 $(\bar{x}(a), \bar{x}(b))$, 通过 φ 的 Clarke 广义梯度和 Ω 的 Clarke 法锥, 横截性条件具有形式

$$(p(a), -p(b)) \in \lambda\partial_C\varphi(\bar{x}(a), \bar{x}(b)) + N_C((\bar{x}(a), \bar{x}(b)); \Omega), \tag{6.119}$$

其中 $\lambda \geqslant 0$. 当 φ 和 Ω 为凸时, 横截性条件 (6.119) 约化为早先由 Rockafellar[1143] 得到的结果. 注意到, 规范形式 $\lambda = 1$ 在平静性假设下成立并且如果 φ 在 $(\bar{x}(a), \bar{x}(b))$ 附近仅为下半连续 (l.s.c.), 那么相应的 (6.119) 式由 $\varphi + \delta(\cdot; \Omega)$ 上图的 Clarke 法锥表示.

横截性条件的更为先进的形式

$$(p(a), -p(b)) \in \lambda\partial\varphi(\bar{x}(a), \bar{x}(b)) + N((\bar{x}(a), \bar{x}(b)); \Omega) \tag{6.120}$$

首先由 Mordukhovich 在 20 世纪 70 年代中期通过基本/极限法锥和次微分建立的: 其中文献 [887] 研究了时间最优控制问题, 文献 [889, 892] 考虑涉及 ODE 控制系统 (6.106) 和微分包含 (6.110) 的最优控制和动态最优化的其他类型的问题; 也请参见文献 [717, 897, 900–902, 904]. 这些结果是由度量逼近的方法得到的, 实际上, 它是文献 [887] 中引入非凸值法锥和次微分的动力; 更多的评论和讨论已在 1.4.5 和 2.6.1 小节中给出.

在西方的文献中, (6.120) 形式的横截性条件起初似乎没有得到适当的关注, 直到 Mordukhovich 在 Montreal 研讨会上所作的报告 (1989 年 2 月) 和 Clarke 的第二部著作 [257](这些条件在书中的脚注中提到, 并归功于 Mordukhovich; 参见小节 1.4.8.) 的出版. 然而, 既使在此之后, 仍然有许多文章 (例如 1.4.8 小节中所列的文章) 继续使用 (6.119) 形式的横截性条件而不是更为先进的 (6.120) 形式.

后来人们逐渐认识到, 在研究到的任何动态最优化的情形, 都有可能证明 (6.120) 这个先进的横截性条件. 这里特别推荐读者参阅文献 [33, 40, 93, 113, 258, 260, 261, 264, 265, 275, 443, 506, 605, 611, 616, 801, 805–807, 845, 847, 878, 880, 914–916, 921, 932, 955, 959, 970, 971, 973, 974, 976, 1021, 1022, 1074–1079, 1080, 1118, 1161, 1162, 1176, 1179, 1211, 1215, 1216, 1233, 1289, 1293–1295, 1372], 这些文献对变分法以及常微分系统最优控制和相应的分布参数版本等各种问题的该条件作了清晰的阐述.

6.5.7 凸值微分包含的广义 Euler-Lagrange 条件

在微分包含动态最优条件的框架下, 文献 [887] 中使用的非凸法锥最初于 20 世纪 80 年代在 Mordukhovich 的文章 [892] 中出现, 他是在凸值、有界和 (关于 x) Lipschitz 的微分包含 (6.110) 的绝对连续轨道上考虑受限于端点约束 (6.118) 的极

小化费用函数 $\varphi(x(a), x(b))$ 的问题. 对于该问题的给定最优解 $\bar{x}(\cdot)$, 文献 [892] 中得到具有形式

$$(\dot{p}(t), \dot{\bar{x}}(t)) \in \operatorname{co}\Big\{ (u,v) \in \mathbb{R}^{2n} \mid (u, p(t)) \in N((\bar{x}(t), v); \operatorname{gph} F(t)),$$
$$v \in M(\bar{x}(t), p(t), t)\Big\}, \quad \text{a.e. } t \in [a,b] \tag{6.121}$$

的动态必要最优性条件, 其中极大点集合 $M(x,p,t)$ 定义为

$$M(x,p,t) := \{v \in F(x,t) \mid \langle p, v\rangle = \mathcal{H}(x,p,t)\},$$

并且当 φ 是局部 Lipschitz 时, 横截性包含 (6.120) 成立. 如果对几乎所有 $t \in [a,b]$, 极大点集合 $M(\bar{x}(t), p(t), t)$ 是一个单点集 (特别地包括当速率集 $F(\bar{x}(t),t)$) 是几乎处处严格凸的情形), 条件 (6.121) 约化为

$$(\dot{p}(t), \dot{\bar{x}}(t)) \in \operatorname{co}\Big\{(u,v) \mid (u, p(t)) \in N\big((\bar{x}(t), \dot{\bar{x}}(t)); \operatorname{gph} F(t)\big)\Big\}, \quad \text{a.e.}. \tag{6.122}$$

值得一提的是, 在文献 [892] 中这些结果是在没有平静性和/或任何其他规范条件下, 利用离散逼近的方法导出的; 关于该技巧的更多讨论请参见 6.5.12 小节.

注意到, 与要求基本法锥完全凸 (因为 $N_C = \operatorname{cl}\operatorname{co} N$) 的 Clarke 的 Euler-Lagrange 条件 (6.115) 相比, 条件 (6.121) 和 (6.122) 只涉及偏凸化, 这样就可避免图像集的 Clarke 法锥的子空间性质所带来的麻烦.

条件 (6.122) 显然蕴涵着 Clarke 形式的 Euler-Lagrange 条件 (6.115); 很容易可以找到 (6.122) 严格好于 (6.115) 的例子. 但是当速率集 $F(x,t)$ 不是严格凸时, (6.115) 和 (6.121) 的关系则并非如此. 事实上, 在 Loewen 与 Rockafellar[805] 中, 有例子表明这两个必要最优性条件一般是相互独立的. 此外, Ioffe[603] 与 Rockafellar[1162] 随后证明了 Mordukhovich 关于凸值微分包含的 Euler-Lagrange 条件 (6.121) 的最初形式等价于 Clarke 的 Hamilton 条件 (6.116) (这两篇文章证明了互补的两个蕴涵关系).

这里推荐给读者 Mordukhovich 另外的文章 [901, 902, 908], 这些文章包含了条件 (6.121), 进而是 (6.122) 在严格凸速率集情形下的发展, 所考虑的各种动态最优化问题涉及凸值 (或松弛) 微分包含, 尤其是具有自由时间, 中间状态约束, 或 Bolza 型泛函等问题. 后来 Smirnov[1215] 发展了文献 [892, 901, 902, 908] 中的离散逼近技巧, 通过将 (非严格) 凸值、Lipschitz、有界和自治微分包含约化为严格凸的情形, 建立了精细化的 Euler-Lagrange 条件 (6.122).

在这个方向上更进一步的结果由 Loewen 与 Rockafellar[805] 对凸值和无界 (6.110) 类型的微分包含得到, 其中将有界包含中 $F(\cdot,t)$ 的标准 Lipschitz 性质替换为无界情形中相应的“可积次 Lipschitz”(integrable sub-Lipschitz) 性质. 该文导

出了 Euler-Lagrange 条件的先进形式 (6.122), 并且强调"我们的方法贯穿了两个简单的主题: 截取和严格凸性", 这说的是他们发展的一个有效的技巧, 可将一般的情形约化为有界和 Lipschitz 的微分包含, 而对这种情形条件 (6.121) 成立并且与精细化 (6.122) 一致. 注意到, 在文献 [805] 发展的技巧中, 集合 $F(x,t)$ 的凸性假设起到了关键的作用. 在 Loewen 和 Rockafellar 随后发表的两篇文章 [806, 807] 中, 将这些结果推广到具有状态约束和自由时间的广义 Bolza 问题中. 值得一提的是, 在文献 [806] 中, (6.112) 式中具有增广实值被积函数/ Lagrange 函数的一般 Bolza 情形, 在适度的"上图连续性"和增长性的假设下, 被约化为一个无界微分包含的 Mayer 问题, 该问题满足文献 [805] 中的"可积次 Lipschitz"性质; 此外, 为阐明该约化的可能性, Mordukhovich[909] 建立的类 Lipschitz 映射的上导数判别法则 (参见定理 4.10) 是一个关键的技术工具.

现在注意到 Euler-Lagrange 包含 (6.122) 可等价地写为上导数形式

$$\dot{p}(t) \in \operatorname{co} D_x^* F(\bar{x}(t), \dot{\bar{x}}(t), t)(-p(t)), \quad \text{a.e.}, \tag{6.123}$$

这实际上是文献 [892] 中引入上导数结构 (作为 F 的伴随映射) 来描述由离散时间和微分包含控制的最优控制问题伴随系统的最初动力. 因为光滑单值映射的上导数约化为伴随 Jacobi 矩阵/函数, 所以关系式 (6.123) 可视为伴随系统 (6.107) 在由微分包含控制的广义控制过程的一个适当推广. 注意到, (6.113) 中 Hamilton 形式的必要最优性条件在非光滑的情形并没有这样的推广. 除了固有的审美价值外, 形式 (6.123) 带有强大的技术成分, 使得能够综合利用上导数分析法则以及 Lipschitz 函数的对偶刻画和相关性质来研究微分包含控制理论中的许多问题, 特别是那些涉及极限过程的问题; 参见本书 6.1 和 6.2 节中给出的主要结果的证明.

6.5.8 非凸值微分包含的广义 Euler-Lagrange 和 Weierstrass-Pontryagin 条件

正如所提到的, 6.5.7 小节中讨论的结果 (以及 6.5.6 小节中先前的版本) 是在没有类平静性假设时在微分包含的速率集 $F(x,t)$ 上所作的凸性假设下导出的. 涉及广义 Euler-Lagrange 条件 (6.123) 的具有端点约束的非凸值 (但 Lipschitz 和有界) 微分包含的必要最优性条件首先由 Mordokhovich[915] 在无任何约束规范条件下建立的. 注意到, Clarke 的完全凸形式的 Euler-Lagrange 条件 (6.115) 先前由 Kaśkosz和 Lojasiewicz[667] 对非凸、有界和 Lipschitz 微分包含的边界轨道得到. 在文献 [915] 中, 读者可以找到非凸微分包含上具有有限非凸被积函数 Bolza 问题 (6.112) 的广义 Euler-Lagrange 条件 (6.123) 的相应版本. 关于自由时间问题可见 Mordukhovich 的另外一篇文章 [916].

Weierstrass-Pontryagin 最大值条件 (6.117) 对凸微分包含并非独立的, 这是因

为它可由上面讨论过的任何形式的 Euler-Lagrange 条件自动得到. 在非凸情形中, 这不再成立, 上述提到的文章 [667, 915] 中并没有建立最大值条件. 但是文献 [915, 注 7.6] 中指出, 在无任何光滑性和/或凸性假设下, 如果能够对具有有限 Lagrange 函数和自由端点 Bolza 问题的强极小点建立经典的 Weierstrass 必要条件, 那么除了精细化的 Euler-Lagrange 条件 (6.123), 所发展的方法能够证明 (6.117). 具体实现首先由 Ioffe 与 Rockafellar[616] 完成, 他们对具有有限 (实值) 被积函数 ϑ 的非凸 Bolza 问题 (6.112) 导出了广义 Euler-Lagrange 条件 (6.123) 的对应版本

$$\dot{p}(t) \in \text{co}\left\{u \in \mathbb{R}^n \mid (u, p(t)) \in \partial\vartheta(\bar{x}(t), \dot{\bar{x}}(t), t)\right\}, \quad \text{a.e.} \tag{6.124}$$

以及经典的 Weierstrass 条件

$$\vartheta(\bar{x}(t), v, t) \geqslant \vartheta(\bar{x}(t), \dot{\bar{x}}(t), t) + \langle p(t), v - \dot{\bar{x}}(t)\rangle, \tag{6.125}$$

它对所有 $v \in \mathbb{R}^n$ 和几乎处处的 t 成立.

基于 Ioffe-Rockafellar 的结果和文献 [915] 中的技巧, 在文献 [914] 中, Mordukhovich 在 F 关于 x 的有界和 Lipschitz 假设下, 对非凸微分包含导出了 Euler-Lagrange 条件 (6.123) 以及 Weierstrass-Pontryagin 最大值条件 (6.117). 该类型的更一般的结果接下来由同时期的文章 Ioffe[604] 和 Vinter 与 Zheng[1294] 建立, 他们分别用不同的方法对无界非凸微分包含导出了广义 Euler-Lagrange 条件 (6.123) 和 Weierstrass-Pontryagin 条件 (6.117) 等必要最优性条件, 其中用到了 Loewen 与 Rockafellar 在文献 [805] 给出的可积下次 Lipschitz 的假设. 有趣的是, 对具有有限 Lagrange 函数的问题, Vinter 与 Zhang[1294] 给出了 Ioffe-Rockafellar 的结果 (6.124) 和 (6.125) 的另外一个证明, 该证明基于将它们约化为具有光滑动态和非光滑端点约束系统的最优控制问题, 并利用具有横截性条件 (6.120) 的最大值原理. 该最大值原理最初由 Mordukhovich 在他的 1976 年的文章 [916] 中得到. 关于具有状态约束和自由时间问题的广义 Euler-Lagrange 和 Weierstrass-Pontryagin 条件的适当形式及其应用, 请读者参阅 Vinter 与 Zheng 随后的文章 [1295, 1296, 1297]. 更进一步, Rampazzo 与 Vinter[118] 将这些结果推广到具有所谓的退化状态约束的非凸微分包含上, 从而对端点可能属于状态约束的边界的问题提供了非退化的必要最优性条件. 对于这之前关于退化控制问题的结果, 也请参见 Arutyunov 与 Aseev [33], Ferreira, Fontes 与 Vinter [443] 以及其中的参考文献.

不久前, Clarke[260, 261] 对非凸和无界微分包含在初始数据的相当弱 (可能是最少) 的假设下, 导出了广义 Euler-Lagrange 形式的必要最优性条件 (6.123), 以及 Weierstrass-Pontryagin 最大值条件 (6.117). 在证明过程中, 他发展了一种涉及光滑变分原理和解耦的强有力的精致方法, 在所作的弱假设下, 将这些条件约化为上面讨论过的已知情形. 文献 [260, 261] 中导出的条件还集成了一种新的分层功能, 其

中的假设和结论是相对于一个预先给定的半径函数来描述的. 另外它们还可以对具有很一般性质的费用被积函数和光滑动态最优控制问题导出所谓的"混合最大值原理".

注意到, 在某些特殊情况下, 可得到变分法和相关问题的极小化非凸和非光滑积分泛函的潜在的更强版本的广义 Euler-Lagrange 条件. 为此请参阅 Ambrosio, Ascenzi 与 Buttazzo [17], Marcelli [845, 846], 以及 Marcelli, Outkine 与 Sytchev [847], 其中对某些具有特殊结构的非凸问题导出了由凸分析的次微分描述的 Euler-Lagrange 条件的某些版本. 这类结果很大程度上依赖于松弛技巧, 尤其涉及 Lyapunov 凸性定理 [822] 及其各种推广和修改.

6.5.9 对偶性与广义 Hamilton 条件的形式

在 6.5.5 和 6.5.7 小节中已讨论过微分包含和广义 Bolza 问题 Euler-Lagrange 条件以前的版本与 Hamilton 最优性条件之间的一些关系. 如前所述, 与经典的光滑和完全凸情形不同, 即使在简单的情形, Clarke 版本的 Euler-Lagrange 条件 (6.115) 和 Hamilton 条件 (6.116) 也不等价. 对于凸值微分包含, Clarke 的 Hamilton 条件事实上等价于早期的 Mordukhovich 版本的 Euler-Lagrange 条件 (6.121). 在没有严格凸性的假设下, 对微分包含情形的广义 Euler-Lagrange 条件 (6.122) (或等价地写为 (6.123)), 和 Bolza 问题情形的 (6.124), 相对应的适当的 Hamilton 条件会是什么样子呢?

该问题首先由 Rockafellar[1162] 在由凸分析中经典公式

$$\vartheta^*(x,p) = \sup_{v\in\mathbb{R}^n}\{\langle p,v\rangle - \vartheta(x,v)\} \tag{6.126}$$

定义的 Legendre-Fenchel 变换 (或共轭对应) 的一般框架下探讨. 由凸分析 [1142], 众所周知, 对任何正常、凸和下半连续 (l.s.c.) 函数 $\vartheta(x,\cdot):\mathbb{R}^n\to\overline{\mathbb{R}}$, 共轭函数 $\vartheta^*(x,\cdot)$ 在 $\mathbb{R}^n$ 上具有相同的性质, 此外, 满足对称二次共轭关系

$$\vartheta(x,v) = \sup_{p\in\mathbb{R}^n}\{\langle p,v\rangle - \vartheta^*(x,p)\}.$$

由 Rockafellar[1162] 陈述并解决的问题是函数 $\vartheta(x,v)$ 和 $\vartheta^*(x,p)$ 关于两个变量的基本次梯度之间的关系. 在某种"上图连续性"的假设下 (当 ϑ 或 ϑ^* 在参考点附近是局部 Lipschitz 时, 该假设自动成立), 文献 [1162] 中建立了下面的凸包关系:

$$\mathrm{co}\Big\{u\in\mathbb{R}^n \mid (u,p)\in\partial\vartheta(x,v)\} = -\,\mathrm{co}\{u\in\mathbb{R}^n \mid (u,v)\in\partial\vartheta^*(x,p)\Big\}. \tag{6.127}$$

对于 $\vartheta(x,v)=\delta((x,v);\mathrm{gph}\,F)$ 的微分包含的情形, 考虑到 (6.126) 和 Hamilton 构造 (6.111), 关系式 (6.127) 约化为

$$\mathrm{co}\Big\{u\in\mathbb{R}^n \mid (u,p)\in N((x,v);\mathrm{gph}\,F)\} = \mathrm{co}\{u\in\mathbb{R}^n \mid (-u,v)\in\partial\mathcal{H}(x,p)\Big\}.$$

文献 [1162] 中给出的 Rockafellar 对偶定理 (6.127) 的证明相当复杂, 它基于有限维空间中凸分析的高级工具, 包括 Moreau-Yosida 的逼近技巧、Wijsman 的上图连续性定理、Attouch 的次梯度收敛定理等.

鉴于关系式 (6.127), 对于凸值微分包含, 等价于广义 Euler-Lagrange 条件 (6.123) 的先进的广义 Hamilton 条件可写为如下形式:

$$\dot{p}(t) \in \mathrm{co}\Big\{u \in \mathbb{R}^n \mid \big(-u, \dot{\bar{x}}(t)\big) \in \partial\mathcal{H}(\bar{x}(t), p(t), t)\Big\}, \quad \text{a.e..} \tag{6.128}$$

对于广义 Bolza 问题 (6.112), 相同形式的广义 Hamilton 条件成立, 其中的 Hamilton 函数对应地定义为 Lagrange 被积函数 $\vartheta(x, p, t)$ 关于速率变量 v 的共轭. 在被积函数 $\vartheta(x, v, t)$ 关于 v 是凸的广义 Bolza 问题的框架下, 满足 Euler-Lagrange 条件 (6.124) 和等价的 Hamilton 形式 (6.128) 的改进假设由 Loewen 与 Rockafellar [806] 给出; 参见 6.5.7 小节中 (6.123) 式之前给出的关于广义 Euler-Lagrange 条件相应的讨论, 由 Rockafellar 的对偶结果 (6.127), 现在这可以等价地关联到 Hamilton 条件 (6.128).

在文献 [604] 中, Ioffe 在比 Rockafellar[1162] 明显弱的条件下建立了 (6.127) 中的包含"$\subset$", 但仍然用了 $\vartheta(x, \cdot)$ 的凸性假设. 利用这个结果, 他证明了 Euler-Lagrange (6.123) 和 Hamilton (6.128) 形式的必要最优性条件对于凸值和无界微分包含成立, 其中将 Loewen 与 Rockafellar [806] 中 $F(\cdot, t)$ 的"可积次 Lipschitz"性质替换为更一般的类 Lipschitz (即 Aubin 的"伪 Lipschitz") 性质. 注意到, Ioffe 的证明清楚地揭示了 Euler-Lagrange 条件 (6.123) 在非光滑最优控制中的枢纽作用, 它在无凸性假设下 (参见 6.5.8 小节) 成立并且直接蕴涵了凸值问题的广义 Hamilton 条件 (6.128). 在该方向上, 即使是对有界和 Lipschitz 微分包含, 非凸问题的 Hamilton 条件 (6.128) 的正确性仍是一个悬而未决的问题.

在如文献 [1162] 中大致相同的假设下, Rockafellar 对偶定理 (6.127) 中包含"$\subset$"的另一个证明后来由 Bessis, Ledyaev 与 Vinter[113] 给出 (也参见 Vinter 的著作 [1289] 中的第 7.6 节). 文献 [113, 1289] 中的证明并不像文献 [604, 1162] 中的那样利用 Moreau-Yosida 逼近, 而是利用更直接和传统 (但相当复杂) 的"邻近分析"技巧.

6.5.10 非光滑最优控制中的其他技巧和结果

值得一提的是, 如 Ioffe [604] 所示的, 6.5.8 小节中讨论的非凸微分包含的先进 Euler-Lagrange 形式明显地蕴涵着具有伴随方程

$$-\dot{p}(t) \in [J_x f(\bar{x}(t), \bar{u}(t), t)]^* p(t), \quad \text{a.e.} \tag{6.129}$$

的参数控制系统 (6.106) 的 Pontryagin 最大值原理的非光滑形式, 其中 $J_x f$ 是 f 关于 x 的 Clarke 广义 Jacobi 函数/矩阵, 这里 Lipschitz 映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ 的广义

Jacobi 矩阵 [252, 255] 定义为经典 Jacobi $m \times n$ 矩阵在点 $x_k \to \bar{x}$ 时的极限集合的凸包; 由基本的 Rademacher 定理 [1114], 该极限集合是非空紧的. 这个涉及伴随方程 (6.129) 的非光滑最大值原理最先由 Clarke[250, 255] 对控制系统 (6.106) 基于逼近过程通过 Ekeland 变分原理直接得到. 也注意到, Ioffe[604] 从微分包含的广义 Euler-Lagrange 形式导出了向量场参数族的最大值原理更为先进的形式, 该形式是由 Kaśkosz与 Lojasiewcz[666] 提出的.

Pontryagin 最大值原理在非光滑控制系统中的最初推广可能是由 Kugushev [722] 发表的, 他利用了某种特定的构造技巧由一列光滑系统来逼近给定的非光滑系统. 然而, 他并没有有效地描述该过程中出现的"次梯度"集合. 关于系统 (6.106) 的非光滑最大值原理的其他早期结果由 Warga[1316, 1317, 1321](从 1973 年末开始) 独立得到, 他利用了一些磨光类型光滑逼近技巧及其映射 $f\colon \mathbb{R}^n \to \mathbb{R}^m$ 的导容概念. 该概念并不是唯一定义的, 在变分分析、最优化和控制的某些情形中, 可以给出比 Clarke 的广义 Jacobi 矩阵更为精确的结果. 然而, 与广义 Jacobi 矩阵相比, 导容的凸包并没有提供更多的信息 (如文献 [1320]), 因此 (6.129) 形式的伴随系统包括了 Warga [1316] 中的情形.

Warga 用来导出必要最优性和可控性条件的方法由 Zhu[1370] 推广到除了标准的有界性和 Lipschitz 条件外, 也满足一些局部选择存在性要求的非凸微分包含, 这些要求被纳入到文献 [1370] 建立的最优性条件中. 这类或相关的条件 (参见 Tuan [1273]) 的明显缺陷是没有任何解析机制来得到所需的选择, 即使是凸包含的情形亦如此. 对显式地涉及逼近或线性类型特定辅助对象的假设或结论, 关于其是否可以构造性地得到验证, 类似的评注对其他一些非光滑最优控制和变分问题的必要最优条件也适用, 这类条件特别见于 Frankowska [464, 465, 468] 以及 Polovinkin 与 Smirnov [1094, 1095]; 也请比照 Ahmed 与 Xiang [6] 中的涉及无穷维微分包含的问题.

注意到, 微分包含必要最优性条件理论的另一个方向主要由俄国学派发展, 其主要目标是, 从涉及系统 (6.106) 类型的光滑最优控制问题的 Pontryagin 最大值原理, 由极限过程导出微分包含的结果. 这样一来, 利用不同的光滑逼近, 一些主要关联到凸值微分包含理论中已知结论的有趣结果由 Arutyunov, Aseev 与 Blagodatskikh [34], Aseev [39, 40, 41] 和 Milyutin [875, 876] 得到; 其中文献 [876] 是 Alexei Alexeevich Milyutin 的最后一篇文章, 是在他去世后才提交和发表的.

在其他发展路线上, 非光滑控制系统 (6.106) 的新结果, 由 de Pinho, Vinter 及其合作者利用微分包含控制系统的适当逼近并借助 Ekeland 变分原理得到. 它们不同于 Clarke 形式的, 具有伴随方程 (6.129) 的非凸最大值原理. 这些结果是通过 Hamilton-Pontryagin 函数 (6.108) (有时也称为没有被极大化的 Hamilton 函数) 的关于变量 (x, p, u) 的联合次梯度来描述的. 这一类型的第一个结果由 de Pinho 与

Vinter [1078] 在"Euler-Lagrange 包含"的名称下 (这似乎与该条件的实质不相符), 对具有端点约束的标准最优控制问题导出的. 后来该名称有了适当的改变, 并且这一类型的结果被称为"涉及没有被极大化 Hamilton 包含 (UHI) 的非凸控制系统的必要最优条件"; 更多的讨论请参见文献 [1076]. 这些作者及其合作者随后的文章 [1074–1080] 包含了 UHI 类型结果在具有状态约束、控制和状态变量上的混合约束、代数-微分约束等最优控制问题上的各种推广. 这一类型的结果对弱极小点特别有效; 也请比照相关的文章 Pales 与 Zeidan [1036]. 与 Clarke 形式的非光滑最大值原理相比, UHI 形式的一个最强的优势 (以及初始动机) 是得到非光滑凸控制问题最优性的必要和充分条件的可能性, 这在 Clarke 形式 (6.129) 中是没有的.

6.5.11 最优控制中的对偶与本原空间方法

注意到, 微分包含最优化所发展的大多数技巧没有用到变分方法及其修改, 而变分方法是经典变分法和处理参数控制系统类型 (6.106) 的最优控制的核心. 这种现象最为重要的技术原因似乎在于, 对微分包含 (6.110), 基于比较给定最优解及小的 (在某种意义下) 局部变分的变分方法不能很好的适应动态约束 $\dot{x} \in F(x)$ 和 $u \in U(x)$ 类型的控制约束的特性, 其中控制区域 $U(x)$ 依赖于状态.

微分包含以及类型 (6.106) 的约束控制系统的必要最优性条件理论的另外一个方法是基于所考虑的整个问题的, 而不仅仅只是其最优解的某种近似/扰动过程. 这可能涉及动态最优化问题的各种近似, 其中包括无右端点约束 (这比较容易处理)、精确惩罚、解耦、离散逼近等; 更多的细节和讨论请参见 Clarke [250, 255], Ioffe [604, 611], Mordukhovich [887, 915], Vinter [1289] 及其中的参考文献.

这种类型的技巧和结果导致了由次微分引导的非光滑最优化和控制的必要条件理论, 它涉及对偶空间中的广义微分结构 (法锥、次微分、上导数). 似乎这一类型的最强的一般结果是由本书中的基本/极限对偶空间结构来表述的, 由于它们内蕴的非凸性, 这不可能由本原空间中的类导数结构 (如切锥和方向导数) 来生成. 这些基本结构使得可以在对偶空间理论的框架下来统一该方向上得到的结果.

作为关联到变分方法及其修改的另一路线、方法和结果是处理本原空间中的各种最优解的变分和扰动, 它涉及各种切向近似, 尤其是控制系统的可达集; 请参见文献 [1102] 中 Pontryagin 最大值原理的证明和后续的文章 Dubovitskii 与 Milyutin [370, 877], Halkin [539, 545], Neustadt [1001, 1002], Warga [1315, 1316] 等. 这一类型的结果称为本原空间理论. 该术语与 Vinter [1289, 225–231 页] 中所采用的不一样.

从所处理的局部极小点、所利用的解析机制和在初始数据上所加的假设的角度来看, 对偶空间和本原空间理论中得到的非光滑最优控制的必要最优性条件一般是相互独立的. 具体地说,

— 本原空间方法探讨的局部极小点依赖于所采用的变分, 而对偶空间方法处理的局部极小点与变分无关.

— 本原空间方法的实施与实现在很大程度上依赖于非线性分析中强有力的工具 (包括开映射和隐函数定理和/或不动点的结果), 而对偶空间方法则避免了这一限制, 即, 在有限维空间中, 利用更为简单的罚函数技巧, 在无穷维的情形, 利用现代变分原理.

— 在对偶空间理论中, 近似/扰动技巧所需的假设要求函数在极小点附近有良好的性质 (例如 Lipschitz 性质和度量正则性), 而本原空间技巧的假设可以只在该点处.

— (光滑和非光滑) 约束最优化 (包括约束最优控制) 的本原空间方法最后需要用凸集分离定理来得到最终以对偶元描述的有效结果 (Lagrange 乘子、伴随轨道等), 而对偶空间方法则不需要凸集分离定理.

在 6.3 节中, 读者可找到一些在本原空间中以经典 PMP 形式和上次微分的推广导出的高等结果. 所得到的结果有关无穷维空间中具有光滑动态和由有限多个实值函数来表述的端点等式和不等式约束的参数化控制系统类型 (6.106). 然而, 这些函数在参考最优点可能仅仅只是 Fréchet可微的, 甚至在该点附近不是连续的 (这只适用于描述端点目标和不等式约束的函数); 请参见下面给出的有关 6.3 节材料的更多评论.

在有限维系统的非光滑最优控制中, 原始类型最一般的结果由 Sussmann 在 20 世纪 90 年代得到; 请参见文献 [1235–1238] 以及其中的文献. 这些工作始于最大值原理的 Lojasiewicz 精细化这个著名结果 [1235], 该结果源于 Lojasiewicz 未发表 (也许未完成) 的文章 [810] 中所阐述的思想. 该精细化包括 PMP 某种形式的证明, 这个证明仅仅假设了 (6.106) 中的速率映射 $f(x,u,t)$ 沿着给定的最优控制 $u=\bar{u}(t)$, 对几乎处处的 t, 关于 x 是局部 Lipschitz 的, 这不同于经典的 PMP 情形, 即对所有 $u \in U$ 和几乎处处的 t, 映射 f 关于 x 是 $\mathcal{C}^1$ 的, 也不同于在 “最少的假设” 下 [250] 的 Clarke 非光滑形式的 PMP 情形, 即对所有 $u \in U$ 和几乎处处的 t, 映射 f 关于 x 是局部 Lipschitz 的. 当 $f(\cdot,\bar{u}(t),t)$ 沿着最优控制 $u=\bar{u}(t)$, 对几乎所有 t, 在点 $\bar{x}(t)$ 可微 (可能不是严格可微) 时, 该结果的 “弱可微” 版本证明了 PMP 的正确性.

Sussmann 通过发展针形变分 (这在经典 PMP 的证明中是关键的) 的某种抽象形式和广义微分的本原空间结构, 在非光滑最优控制中, 证明了这些结果及其深远推广. 在最近的文章 [38] 中, Arutyunov 与 Vinter 提供了 PMP 的 Lojasiewicz 精细化中 “弱可微” 版本的一个简化证明. 该证明基于所谓的 “内部有限逼近”, 它涉及在轨道上没有违反端点约束的参考最优控制 $\bar{u}(\cdot)$ 的特殊针型变分. 这个有限逼近方案的思想可追溯到 Tikhomirov 的文章 [7], 在那里, 该思想在光滑最优控制的

经典 PMP 中用到. 该方向上的更进一步的结果由 Shvartsman[1209] 对具有状态约束的非光滑控制系统导出.

6.5.12　离散逼近方法

6.1 节致力于用离散逼近的方法, 在无穷维空间中, 全面研究动态最优化问题. 虽然发展这种方法的主要目的是把它作为导出由非凸微分/发展包含控制的动态过程的广义 Euler-Lagrange 类型 (6.123) 的必要最优性条件的工具, 但是这里也给出了一些数值的结果, 它们是关于涉及和没有涉及最优化的发展包含离散逼近的适定性和收敛性问题. 除了作者最近的文章 [932] 外, 似乎无穷维发展包含的必要最优性条件或是这些过程的离散逼近均没有在以前的文献中考虑过, 本书中所得到的一些结果在文献 [932] 中已提到过. 然而它们是有限维的系列发展; 请参见下面的讨论.

研究连续时间系统的离散逼近方法可追溯到 Euler [411], 在一维变分法中, 对极小化积分泛函, 他发展了该方法, 建立了著名的一阶必要条件 (现称为 Euler 或 Euler-Lagrange 方程). 注意到, Euler 将极小化下的积分视为无穷和, 不是利用极限过程, 而是把微分 (通过几何图形) 解释成沿着极小曲线的无穷小变化, 从而可类比于"折线", 即有限差分. Euler 对变分法中的"一般"(在那个时候的) 问题以一个方程的形式导出了必要最优性条件, 这标志着一个重大的理论成就, 它综合了出现在早期研究工作的许多特殊情形和例子. 值得一提的是, 将曲线替换为折线的近似思想是在变分法发展之初, 由 Leibniz[757] 在他解决速降线问题中 (相当含糊地) 部分地利用到.

从那时起, Euler 的有限差分方法及其修改已被广泛地应用到微分系统的动态最优化和数值分析的各个领域, 其中特别要强调在数值分析方面的应用, 这在计算机时代变得更加重要. 大量的文献对离散逼近的不同方面及其应用进行研究. 对涉及动态最优化和控制系统的代表性参考文献, 这里推荐给读者文献 [28, 98, 184, 185, 220, 221, 298, 299, 302, 303, 338, 343–349, 353, 354, 357–359, 367, 407, 425, 488, 520, 535, 542, 702, 721, 760, 828, 831, 832, 890, 892, 900–902, 908, 915, 916, 941, 959, 973, 974, 976, 1012, 1061, 1062, 1086, 1107, 1109, 1175, 1215, 1216, 1280, 1282–1284, 1301, 1333, 1379] 及所录文献.

在 6.1 节中, 我们将 Mordukhovich[915] 以前在有限维空间中发展的有关微分包含离散逼近方法的基本结构和结果推广到非凸发展/微分包含的一般无穷维情形; 也请参见文献 [890, 892, 901, 902, 908, 1107, 1109, 1215, 1216] 以及下面对以前关于有限维空间中凸图和凸值微分包含这个方向上所做工作的评论.

涉及微分包含变分问题必要最优性条件的离散逼近方法的基本思想和基本策略, 包括下面三个主要组成部分:

(Ⅰ) 用一列适定离散时间最优化问题序列来替换/逼近初始连续时间变分问题, 使得离散时间最优化问题的最优解序列, 在某种适当的意义下, 收敛于初始问题的某个 (或给定的) 最优解;

(Ⅱ) 将动态最优化离散时间问题约化为数学规划的约束问题, 它们是内蕴非光滑的, 然后利用适当的具有良好分析法则的广义微分工具来导出动态最优化离散时间问题的必要最优性条件;

(III) 由离散逼近的必要条件, 通过取极限并且利用所得到的离散逼近过程的收敛性/稳定性结果和确保所需的伴随轨道收敛的广义微分结构的相应性质, 建立鲁棒/点基必要最优条件.

在 Mordukhovich 的文章 [915] 中, 离散逼近方案是应用在有限维空间中由非凸微分包含控制的广义 Bolza 问题上; 由此非凸问题的先进类型的广义Euler-Lagrange 条件 (6.123) 首先在此文章中建立. 对于无穷维发展包含, 要实现上面三个步骤中的每一个, 都需要一些额外的工作, 其中大部分明显地不同于有限维空间中的情形.

6.5.13　发展包含的离散逼近

Deimling[314] 和 Tolstonogov[1258] 的著作中给出了无穷维空间中微分包含 (6.1) (通常称为发展包含) 理论的主要方面; 更多的著作是关于有限维情形的; 例如 Aubin 与 Cellina [50] 和 Filippov [5[illegible]] 以及其中的参考文献. 本书的定[illegible]遵循 Deimling [314] 中关于 Banach 空间中微分/发展包含解的定义. 注意到, 这有别于有限维 Carathéodory解的定义 (在微分方程的情形这可追溯到文献 [222]); 根据 Bochner 积分, 这里为使 Newton-Leibniz 公式成立, 就需要附加额外的要求. 该公式对于具有无穷维值的绝对连续映射并不自动成立. 另一方面, 有一个一类 Banach 空间的精确刻画, 其中满足 Newton-Leibniz 公式等价于绝对连续性, 此即是具有 Radon-Nikodým性质 (RNP) 的空间, 更多的细节可在 Bourgin [169] 以及 Diestel 与 Uhl [334] 等经典专著中找到. Radon-Nikodým性质 (RNP) 是泛函分析中的基本性质; 尤其, 对偶空间 X^* 的 RNP 等价于 X 的 Asplund 性质. 这导致了著名的 Asplund 空间类在本书中另外一个方向的应用.

6.1.1 小节的主要结果, 定理 6.4, 构造性地证明了在任意 Banach 空间 X 中, 对任意给定的 Lipschitz 微分包含 (6.1) 的可行轨道, 它可由通过 Euler 标准方法得到的有限差分包含 (6.3) 的增广轨道强逼近 (在 Sobolev 空间 $W^{1,2}([a,b];X)$ 中的范数下, 这保证了关于速率的几乎处处逐点收敛). 这个结果是 Mordukhovich[195, 定理 3.1] 的无穷维版本 (证明只需稍作改动), 这个定理是 Mordukhovich[901, 902] 以前的构造和结果以及 Smirnov 的文章 [1215] 中结果的推广; 也请参见文献 [1216]. 除了其独立的意义和数值分析上的价值 (提供了逼近一般微分包含可行解集的有效程序而不用顾及到最优化), 这个定理建立了构造连续时间发展系统变分问题的

适定离散逼近的基础.

注意到, 在定理 6.4 中, 在速率集 $F(x,t)$ 上没有作任何凸性假设, 并且根据 (6.10) 中速率的投影实现了邻近算法. 这把速率方法与关于 (凸图或凸值) 微分包含逼近的传统结果区别开来, 这些结果涉及状态向量的投影并且仅仅只是确保轨道的 $\mathcal{C}([a,b];\mathbb{R}^n)$ 范数收敛; 参见例如 Pshenichnyi [1107, 1109] 以及综述文章 Dontchev 与 Lempio [359] 和 Lempio 与 Veliov [761]. 需要强调的是, $\mathcal{C}([a,b];\mathbb{R}^n)$ 范数收敛性不能处理非凸包含 (因为轨道的一致收敛性对应着导数的弱收敛, 由之就需要接下来的基于 Mazur 弱闭包定理的凸化); 因此, 能够达成导数/速率的几乎处处逐点收敛性对非凸问题必要最优条件的建立是举足轻重的.

下面提一下关于离散逼近的收敛性的两个近期进展, 这对应于 6.5.12 小节中罗列的方面 (i). 在文献 [343] 中, Donchev 推广了前面提到的文章 [915] 中的一些逼近和收敛结果, 其中考虑的有限维微分包含的右边的映射 $F(x,t)$ 对 x 满足 Kamke 条件, 即其标准的 Lipschitz 模换成了 Kamke 类型的函数. 该条件在所有连续多值函数 $F(\cdot,t)$ 组成的类上是“一般”(generic) 的 (在 Baire 的意义下). 另外一个发展见 Mordukhovich 与 Pennanen [941], 该文在特定的凸性和 Lipschitz 假设下建立了广义 Bolza 框架下离散逼近的上图收敛性.

6.5.14 中间局部极小点

6.1.2 小节中开始在 Banach 空间中研究约束微分/发展包含的 Bolza 问题, 主要是沿着 Mordukhovich[915] 在有限维空间中发展的步骤, 其中也有一些在无穷维空间中的显著改变, 见下面的评注. 注意到, 与具有增广实值函数 φ 和 ϑ 并隐含地纳入了端点和动态约束的广义 Bolza 问题 (6.13) 相比, 这里处理的这类约束是显式形式的, 这是因为在 6.1 节所得的结果中, 关于 φ 和 ϑ 的连续性和 Lipschitz 假设事实上排除了这些函数取无穷值的可能性.

该小节研究的主要精力放在秩为 $p\in[0,\infty)$ 的中间局部极小点 (i.l.m.; 见定义 6.7) 及松弛的中间局部极小点 (r.i.l.m.; 见定义 6.12) 的概念上. 这两个概念均由 Mordukhovich[915] 引入, 后来由 Ioffe 与 Rockafellar[616], Ioffe[604], Vinter 与 Woodford[1293], Woodford[1331], Vinter 与 Zheng [1294, 1295, 1289], Vinter[1289], 以及 Clarke[260, 261] 在各种动态最优问题中进行研究, 大多数是 $p=1$ 的情形, 并被称为 $W^{1,1}$ 局部极小点.

对于变分问题, 中间局部极小点处于经典的弱和强极小点之间的中间位置; 这是文献 [915] 中采用中间局部极小点名称的原因. 例 6.8~ 例 6.10 表明了这三种主要类型的局部极小点是不同的, 即使是在比较简单的, 特别是涉及凸值、有界和 Lipschitz 微分包含的动态最优化问题中亦如此. 关于弱和强极小点区别的例 6.8 是经典的, 可追溯到 Weierstrass [1326]. 关于弱和中间极小点区别的例 6.9 的简化版

本见 [915], 该例以及例 6.10 的完整版本取自 Vinter 与 Woodford [1293]. 该文和 Woodford 的博士论文 [1331] 也包含了其他例子来阐述这些局部极小点概念之间的区别, 尤其是有限维空间中凸和无界微分包含的各种秩的中间极小点之间的区别.

6.5.15　松弛稳定性和隐含凸性

6.1.2 小节的其余部分给出了微分包含的松弛 Bolza 问题的结构, 以及和松弛稳定性相关的定义和讨论. 恰当的松弛 (或扩张、推广、正则化) 的思想在现代变分理论中发挥了显著作用. 泛泛而言, 这可追溯到 Hilbert[567] 在他著名的第二十个问题中的表述"在变分法中, 只要对解这个词有恰当的解释, 则每个问题都有解."

对该思想的完整理解始于 20 世纪 30 年代 Bogolyubov[121] 和 Young[1349, 1350] 对于一维变分问题的独立工作. 他们指出, 可以由关于速率的某种凸化来得到变分问题的适当扩张, 这确保了广义最优解及其由"普通曲线"逼近的存在性. 在最优控制中, 该思想由 Gamkrelidge[495] 和 Warga[1313] 独立发展; 术语"松弛"在文献 [1313] 中最先引入. 对于类似的问题, 另外一个广泛用到的术语是"Young 测度". 对于各种松弛结果及其在变分法、最优控制和相关问题中的应用, 请参阅文献 [3, 4, 25, 31, 50, 75, 212, 213, 231, 232, 235, 237, 246, 255, 308, 362, 401, 432, 450, 497, 527, 617, 618, 682, 704, 821, 823, 863, 886, 888, 901, 915, 1020, 1049, 1082, 1173, 1174, 1176, 1177, 1258, 1259, 1277, 1315, 1323, 1351] 及其文献.

本书遵循文献 [915] 中发展的涉及有限维微分包含 Bolza 问题的结构并利用松弛过程. 这里利用松弛过程并不是为了确保广义解的存在, 而是可以描述离散逼近问题最优解以及极小化泛函的值的极限点. 在此过程中, (6.19) 中松弛稳定性的概念起到了重要的作用. 该性质是连续时间控制系统和微分包含所固有的, 并关联到它们的隐含凸性; 有关的更多讨论和充分条件请参见 6.1.2 小节以及其中的参考文献. 特别注意到, 定理 6.11 中的逼近性质取自 De Blasi, Pianigani 与 Tolstonogov 最近的文章 [308], 该逼近性质是无穷维微分包含一般 Bolza 问题框架中隐含凸性的表现. 也注意到, 在深层的意义下, 隐含凸性可追溯到关于无原子向量测度值域凸性的经典 Lyapunov 定理 [822] 和关于集值积分的 Aumann 定理 [55]; 对于无穷维空间中相应的结果请参见 Arkin 与 Levin [25] 以及 Diestel 与 Uhl [334]. 下面指出一些关于隐含凸性的其他显著表现:

— 非凸规划中"对偶间隙"的估计, 它是由 Ekeland[398] 发现并由 Aubin 与 Ekeland[51] 发展. 这些发展与数理经济学中经典的 Shapley-Folkman 定理密切相关; 更多的细节和讨论请参见 Ekeland 与 Temann 的著作 [401].

— "小球的非线性映像"的凸性最近由 Polyak[1098, 1100] 发现, 他得到了该现象在最优化、控制和相关领域中的各种应用; 对于更进一步的发展也请参见 Bobylev, Emel'yanov与 Korovin [120].

6.5.16 离散逼近的收敛性

6.1.1 小节中的主要精力放在不涉及最优化的微分/发展包含的有限差分逼近上, 6.1.3 小节的结果则涉及整个 Bolza 变分问题的逼近. 目标是构造原始 Bolza 问题 (P) 的适定离散逼近, 即离散时间动态最优化问题序列, 使得离散逼近的最优解, 在某种事先指定的意义下, 收敛到连续时间问题的最优解. 事实上, 该小节给出了适定性/稳定性结果, 从而阐明了下面两种类型的离散逼近收敛:

(I) 值收敛. 它确保构造性建立的离散逼近问题费用泛函的最优值收敛到原始问题费用泛函的最优值 (下确界), 这里并没有假设原始问题最优解的存在性.

(II) 强收敛. 即离散时间问题的最优解强收敛到原始问题的给定最优解; 对于分段线性扩张的离散轨道, 强收敛理解为 $W^{1,2}$ 范数下的收敛.

注意到, 类型 (II) 的结果明确涉及原始问题的给定最优解 (实际上是中间极小点). 它们不再是构造性的 (从数值的角度来看), 但可以由之阐明利用离散逼近 (而不是利用变分的方法, 在该框架下, 变分的方法不再适用) 来导出连续时间问题的必要最优性条件的方法. 6.1.3 小节中得到的类型 (II) 的收敛性结果的主要目的在于导出本书 6.1 节中的必要最优性条件 (也请比照 7.1 节中关于泛函微分控制系统相应的必要最优性条件); 与证明 (I) 中的值收敛所需的假设条件相比较, 它们通常要求较为适中的假设.

类型 (I) 的结果传统地关联到最优控制中的计算方法; 基于连续时间控制问题的有限差分序列逼近, 它们阐明了"直接的"数值技巧. 如果控制系统中的状态向量是有限维的, 那么最优逼近问题可约化为有限维的数学规划问题. 除了 6.1.3 小节给出的结果外, 作者不熟悉该方向上有关无穷维微分包含的任何结果, 即使是参数化控制形式 (6.106).

关于标准控制系统 (6.106) 值收敛的第一个结果可能由 Budak, Berkovich 与 Solovieva[184] 以及 Culum[302] 于 20 世纪 60 年代末在相当苛刻的假设下得到; 有关早期的进展也请参见文献 [185, 303, 407]. 之后, Mordukhovich[890] 建立了, 如果状态/端点约束的适当扰动与离散化步长相一致, 那么离散逼近的值收敛和涉及参数系统 (6.106) 的一般控制问题的松弛稳定性是等价的. 这些结果在文献 [899, 901, 902] 中被推广到 Lipschitz 微分包含的情形; 也请比照 Dontchev [349] 以及 Dontchev 与 Zolezzi [367] 中相关的结果. 对于特殊结构系统, 关于费用函数, 以及关于控制和轨道的收敛速度的有效估计由 Hager[535], Malanowski[831], Dontchev[347], Dontchev 与 Hager[355], Veliov[1284], 以及其他作者导出; 更多的细节和参考文献请参见文献 [352, 359, 761] 中的综述.

定理 6.14 似乎是新的结果, 即使是在有限维的情形亦如此, 它发展了 Mordukhovich[890, 899, 901] 中对有限维微分包含所发展的相对应的方法和结果. 注意到,

与有限维的情形相比较, 该定理和相关定理 6.13 的证明在技巧上要复杂得多, 除了用到其他的结果, 该证明还基于确保 $L^1([a,b];X)$ 中序列弱紧性的基本 Dunford 定理, 其中要求空间 X 和 X^* 都满足 Radon-Nikodým性质, 这在 X 和 X^* 都是 Asplund 空间时是成立的. 当然不能忘了, 即使不涉及 RNP, Asplund 结构在本书中发展的广义微分理论中已经起到了至关重要的作用.

对以离散逼近方法作为工具导出连续时间系统 (即是在理论上, 而非数值上的应用) 必要最优性条件, 定理 6.13 是这里实际用到的结果, 它是 Mordukhovich [915] 中定理 3.3 的无穷维推广和修改. 这两个结果之间的区别 (即使在有限维空间中) 涉及逼近原积分泛函的方式: 这里采用构造 (6.20) 而不是如文献 [915] 的简化的 (6.28). 经过这个修改之后, 就可以处理关于 t 的可测被积函数, 这对 6.2 节中的应用是重要的, 因为其中的被积函数必须是可测的.

注意到, 逼近给定中间极小点 $\bar{x}(\cdot)$ 的导数的 (6.20) 和 (6.28) 中的最后一项是很重要的. 由于这一项的存在并利用定理 6.4 中的逼近结果, 即可以建立离散逼近问题最优解到原问题的给定局部极小点 (在 $W^{1,2}([a,b];X)$ 的范数下) 的强收敛性; 进一步, 通过对离散时间问题的必要条件取极限, 就导出了连续时间问题 (6.123) 类型的必要条件. 除了文献 [915] 外, 该逼近项先前由 Smirnov[1215] 对有限维空间中涉及凸值、有界和自治微分包含的 Mayer 问题使用过 (也参见他的著作 [1216]). 在凸值或甚至图凸微分包含的 Mayer 框架中, 这之前也有尝试利用离散逼近来导出必要最优性条件, 但一般仅仅能够确保增广离散轨道一致收敛到 $\bar{x}(\cdot)$, 其中利用 "状态类型" 的逼近项

$$\sum_{j=0}^{N-1} \|x_N(t_j) - \bar{x}(t_j)\|^2,$$

它不涉及导数 $\dot{\bar{x}}(\cdot)$; 请比照 Halkin [542], Pshenichnyi [1107,1109] 和 Mordukhovich [892, 901, 902].

6.5.17 离散逼近的必要最优性条件

在建立了上面讨论过的所需要的离散逼近强收敛/稳定性之后, 为了建立约束微分包含必要最优性条件, 实现离散逼近方法策略的第二步是导出 6.1.3 小节中阐述的离散时间问题的必要条件. 这里考虑了以下两种形式的离散逼近问题:

—"积分" 形式问题 (P_N), 它涉及极小化费用泛函 (6.20), 约束为 (6.3), (6.21)~(6.23) 以及

—"简化" 形式问题 $(\overline{P}_N)$, 它在相同的约束下极小化另外一个费用泛函 (6.28).

如上所述, 泛函 (6.20) 和 (6.28) 两者之间的唯一区别涉及逼近原连续时间 Bolza 问题 (P) 中积分泛函的不同方式: (6.20) 中的积分类型允许考虑 (6.13) 中的可测被积函数 $\vartheta(x,v,\cdot)$, 但 (6.28) 中的和式/简化类型要求关于 $\vartheta(x,v,\cdot)$ 的几乎

处处连续性假设. 考虑简化类型逼近的原因是, 与如 (6.20) 中积分逼近情形的自反性和可分性要求相比, 在更为一般的 Asplund 状态空间 X 情形中, (6.28) 中的和式形式可能得到离散时间问题, 甚至连续时间问题的必要最优性条件. 这是因为对有限和的次微分有更为成熟的分析法则, 而对积分泛函却非如此; 见下面讨论.

在 6.1.4 小节中, 在无穷维空间中, 导出了离散时间动态最优化问题 (P_N) 和 $(\overline{P}_N)$ (以及它们具有较少结构的相应问题 (DP), 称为离散时间包含 Bolza 问题) 的必要最优性条件. 对于具有固定步长的离散系统, 这些问题当然对具有固定步长的离散系统有独立的意义, 这对许多应用是重要的, 特别是在经济动态模型中的应用; 例如参见 Dyukalov [379] 以及 Dzalilov, Ivanov 与 Rubinov [380]. 更进一步, 由 6.1.3 小节中的收敛结果, 它们的必要最优性条件为所考虑的连续时间 Bolza 问题提供了次最优性条件. 当然, 这里主要的兴趣是, (P_N) 和 $(\overline{P}_N)$ 的这类必要最优性条件便于取极限来建立涉及无穷维微分包含 Bolza 问题的必要最优性条件.

6.1.4 小节中考虑的离散时间动态最优化问题可约化为 (6.29) 给出的约束的数学规划 (MP) 形式. 以这种方式出现的问题 (MP) 有两个特征, 有别于数学规划中其他类型的约束问题:

(a) 它们涉及有限多个几何约束, 当离散逼近步长趋于零时, 其数目趋于无穷. 值得一提的是, 这些几何约束是图像类型的, 它们是由离散化包含生成的. 这类约束的存在使 (MP) 问题 (6.29) 是内蕴非光滑的, 即使对 (6.29) 和生成的问题 (P_N), $(\overline{P}_N)$ 和 (P) 中的光滑泛函数据亦如此.

(b) 如果原状态空间 X 是无穷维的, 那么 (MP) 问题 (6.29) 不可避免地包括等式类型 $f(x) = 0$ 的算子约束, 其中 f 的值域空间不可能是有限维的. 如所周知, 这类约束是最优化中最为困难的, 即使对光滑映射 f 亦如此; 实际上, 所考虑的在离散时间问题中的应用便是如此.

类型 (6.29) 的数学规划问题的必要最优性条件理论见第 5 章, 其中的必要条件是以基本/极限广义微分构造描述的. 这类涉及推广 Lagrange 乘子问题的主要条件总结于命题 6.16 中, 其中 (6.29) 中的有限多个几何约束, 在 Asplund 空间的框架下, 通过基本法锥的交法则和相应的 SNC 分析法则结果纳入. 利用 (MP) 的这些最优性条件和对基本法向量及次微分所发展的确切/点基分析法则, 可得到广义 Euler-Lagrange 形式的定理 6.17 中由差分包含控制的离散 Bolza 问题 (DP) 的必要最优性条件. 注意到, 该结果并没有在离散速率集 $F_j(x)$ 上作任何凸性和/或 Lipschitz 假设. 定理 6.17 中得到的条件是 Mordukhovich [915, 定理 5.2] 中有限维条件相对应的 Asplund 空间中的形式, 其中有关于 SNC 的特定条件, 对无穷维情形是需要的.

定理 6.17 中得到的离散 Bolza 问题 (DP) 的点基必要最优性条件, 就其自身而言, 是重要的. 并且, 更进一步, 为导出无穷维空间中连续时间问题广义 Euler-

Lagrange 类型 (6.123) 的必要最优性条件提供了足够的基础; 更多的细节请参见文献 [915]. 然而, 在无穷维空间中却不完全如此, 此时该方法的实现要求额外的 SNC 假设, 从而能够确保离散逼近中的点基必要最优性条件成立, 并且当 $N \to \infty$ 时, 能够对这些离散逼近的必要最优性条件取极限. 对离散时间问题的近似/模糊必要条件, 而不是定理 6.17 中的点基必要条件, 这些额外假设可以避免. 这样的近似最优条件是在定理 6.19 和定理 6.20 中, 对离散逼近问题 $(\overline{P}_N)$ 和 (P_N) 分别得到的.

这里提到的近似最优性条件的证明相当复杂, 除了其他要求之外, 还要求利用模糊分析法则以及由 Mordukhovich 和 Shao[946] 建立的度量正则性的邻域上导数刻画. 也注意引理 6.18 的显著作用, 它将积分号下取 (次) 微分的经典 Leibniz 法则推广到基本次梯度的情形. 这是定理 6.20 证明的一个辅助结果, 这样, 在离散逼近类型 (P_N) 下, 可处理 (P) 中的可和被积函数; 当然该法则也有其独立的意义. 它的证明利用了 Lyapunov-Aumann 凸性定理的无穷维推广和 Clarke 次梯度的相应法则 [255, 定理2.7.2], 而该法则进一步是严重基于由 Ioffe 和 Levin[612] 对凸分析中次梯度建立的 Leibniz 法则的广义形式.

6.5.18 由离散逼近取极限

在 6.1.5 小节中, 完成了导出微分包含原 Bolza 问题 (P) 的必要最优性条件的离散逼近方法的第三步 (在 6.5.12 小节中记为 (iii)). 这一步的主要目的是证明从所得到的适定离散逼近问题 (P_N) 和 $(\overline{P}_N)$ 的必要条件中取极限, 并且有效地描述由该过程得到的连续时间问题的必要最优性条件. 正如在该小节看到的, 所得到的条件是定理 6.21 和定理 6.22 中建立的对 (P) 的松弛中间局部极小点的广义 Euler-Lagrange 类型的条件.

6.1.5 小节中的这些主要结果, 在所作的假设及 (6.44) 和 (6.47) 中广义 Euler-Lagrange 包含的阐述方面, 相互之间有所不同. 这些差异来自 6.1.4 小节中关于两种类型的离散逼近问题 $(\overline{P}_N)$ 和 (P_N) 相应的结果, 以及在这些问题的必要最优性条件中取极限所需的额外要求.

基于简化离散逼近 $(\overline{P}_N)$ 上的极限过程, 定理 6.21 是 Mordukhovich[915, 定理6.1] 的无穷维推广, 它涉及 (6.44) 中的广义法锥. 在文献 [915] 类似的情形中, 基本法锥的利用是由某种技巧性的假设来支持, 它确保了定义 5.69 中阐述的正则连续性, 该性质在此定义后有讨论. 定理 6.22 的结果是新的, 即便是在有限维的情形亦如此.

在定理 6.21 和定理 6.22 的证明中, 从离散时间必要最优性条件中取极限的主要顾虑之一是证明伴随轨道及其导数的适当收敛. 为了建立所需的收敛, 这里利用集值映射 Lipschitz 性质的对偶上导数刻画, 它在本书中经常用到; 在由 Lipschitz 映射描述的动态最优化问题中, 该准则在完成关联到离散时间和连续时间包含的伴随系统的极限过程中起到了关键的作用.

无穷维微分包含必要最优性条件与其相对应的有限维情形的必要最优性条件的区别主要在于, 定理 6.21 和定理 6.22 中关于约束/目标集 Ω 的 SNC (实际上是强 PSNC) 假设. 这类假设对无穷维发展系统的最优控制问题是至关重要的. 特别地, 众所周知, 在 Hilbert 空间中, 即便是对具有单点目标集 $\Omega = \{x_1\}$, 由一维热传导方程控制的简单最优化控制问题, Pontryagin 最大值原理也不成立. 在无穷维空间中, 该单点集永远不是 PSNC 的. 这种类型的第一个例子由 Y. Egorov[393] 给出. 对于涉及有限余维数性质 (对凸集合来说这等价于 SNC 性质; 见注 6.25) 的更多讨论, 读者也可以参阅 Fattorini[432] 以及 Li 与 Yong [789] 的著作. 为此需要强调推论 6.24 中的结果, 它证明了由发展包含控制的且在约束集 Ω 上没有显式 (其实是隐含) 的 SNC/PSNC 假设的 Bolza 问题的广义 Euler-Lagrange 条件, 其中约束集是由有限多个等式和不等式通过 Lipschitz 函数给出的.

最后推荐 Mordukhovich 和 D.Wang 最近的文章 [970, 971] 给读者, 其中对于由半线性无界发展包含控制的最优控制问题, 导出了一些与上面相对应的结果. 半线性无界发展包含对抛物型偏微分方程的建模特别方便; 请参见注 6.26.

6.5.19 无松弛的 Euler-Lagrange 和最大值条件

如前所述, 在 6.1 节中, 由离散逼近方法建立的广义 Euler-Lagrange 条件适用于由无穷维微分包含控制的 Bolza 问题的松弛中间局部极小点. 6.2 节的主要目标是, 基于 6.1 节中得到的条件并且涉及额外的变分技巧, 对无任何松弛的非凸微分包含导出 Euler-Lagrange 类型的精细化结果, 以及进一步的 Weierstrass-Pontryagin 最大值条件. 为了简单起见, 定理 6.27 只考虑具有固定左端点和在右端点上作任意几何约束的 Mayer 型问题 (P_M), 这在无穷维空间中是新的结果; 之前的有限维形式在 6.5.8 小节中讨论过.

如 6.1 节中那样, 在有限维和无穷维情形得到的必要条件之间的主要区别在于, 在无穷维中 SNC 的要求是不可避免的. 另一方面, 定理 6.27 中所作的技术性假设与定理 6.21 和定理 6.22 中所作的不同. 也注意到, 与 6.1 节中涉及 Lipschitz 费用和约束函数的主要结果相比, 定理 6.27 中的横截性条件 (6.51) 和 (6.52) 具有更一般的形式.

除了利用变分分析中相当微妙的微积分和收敛结果外, 无松弛非凸问题的中间局部极小点的 Euler-Lagrange 条件 (6.49) 的证明是基于两个扰动/逼近过程, 从而将原问题 (P_M) 约化为无约束 (但仍是非光滑和非凸的) Bolza 问题 (6.55). 该无约束 Bolza 问题具有关于状态和速率变量是 Lipschitz 的并且关于 t 是可测的有限值数据. 因为无约束 Bolza 问题的任何中间局部极小点自动是松弛问题的一个中间局部极小点, 所以可以通过离散逼近由定理 6.22 中得到的必要最优性条件来处理.

上述的第一个扰动技巧可认为是由 Mordukhovich[887] 最初在证明具有光滑动

态和非光滑端点约束的有限维控制问题的最大值原理时发展的度量逼近方法, 其中的问题约化为自由端点问题. 第二个技巧涉及 Ekeland 变分原理和动态约束的罚方法, 这可追溯到 Clarke[251] 关于有限维中非光滑控制系统的 Hamilton 函数和最大值条件. 定理 6.25 证明中的论断是 Kaśkosz和 Lojasiewicz[667] 建立的关于强极小点 (或边界轨道) 相应结果的无穷维推广. 注意到, 1.3.3 小节中关于距离函数在集合里的点和集合外的点的广义微分结果的重要性, 它们用来处理逼近问题且对 Lipschitz 性质的上导数准则起到了至关重要的作用, 这样就可以在导出定理 6.27 中的广义 Euler-Lagrange 和横截性包含时达成收敛过程.

最大值条件 (6.50) 是广义 Euler-Lagrange 条件 (6.49) 在非凸情形的补充, 6.2.1 小节中只概述而不是完整地给出其证明, 这是因为该证明虽然在技巧上相当复杂, 但密切遵循 Vinter 和 Zheng[1294](也参见 Vinter 的著作 [1289, 定理 7.4.1]) 对有限维微分包含情形所给出的证明; 读者可自行验证所有细节. 注意到, 该证明基于将微分包含的一般 Mayer 问题约化为具有光滑动态和非光滑端点约束的最优控制问题, 这首先由 Mordukhovich[887] 通过非凸/极限法锥来处理; 无穷维空间情形相关的控制问题和技巧请参见 6.3 节. 在 Euler-Lagrange 框架 (6.49) 中, 由 Ioffe[598] 和 Clarke[261] 给出的其他证明似乎仅适用于有限维状态空间.

6.5.20 微分包含最优控制中相关的论题和结果

利用本书发展的变分方法可以得到定理 6.27 在各种情形中的推广和相应的结果, 其中部分已在 6.2.2 小节中讨论过, 特别地, 包括上次微分条件和多目标控制问题; 对于涉及有限维控制系统的多目标动态最优化中相关的进展, 也请比照 Zhu [1372], Bellaassali 与 Jourani [93], 以及 Eisehart [395]. 然而, Hamilton 类型的必要最优性条件以及关于微分包含局部可控性的结果似乎要求状态空间是有限维的; 请参见注 6.32 和注 6.33 中更多的细节和讨论.

6.2.2 小节末给出的例子表明了所得到的关于微分包含的结果的一些特性和它们之间的关系. 例 6.34 证实了偏凸化对广义类型的 Euler-Lagrange 和 Hamilton 最优性条件的正确性是必不可少的, 这归功于 Shvartsman (个人交流). 取自 Loewen 与 Rockafellar [805] 的例 6.35 表明, 只涉及偏凸化的广义 Euler-Lagrange 条件严格好于 Clarke 完全凸化形式的 Hamilton 条件, 即使对具有凸速率的 Lipschitz 控制系统亦如此. 最后, 由 Ioffe [604] 给出的例 6.36 表明, 偏凸化 Hamilton 条件在相当广泛的情形也严格优于完全凸化的 Hamilton 形式, 其中偏凸化 Hamilton 条件可能不等价于其相应的 Euler-Lagrange 条件.

6.5.21 基于增量方法的本原空间方法

6.3 节在更传统的参数化框架 (6.61) 中考虑涉及无穷维动态的最优控制问题,

更进一步, 在本节中, 在速率函数 f 上关于状态变量 x 作连续可微性/光滑性假设. 然而, 对于由非光滑发展包含控制的动态最优化问题, 6.3 节中给出的结果至少在下面的主要方面与 6.1 节和 6.2 节中得到的结果不同:

— 在问题中, 状态空间是任意的 Banach 空间, 没有额外的几何假设;

— 目标和 (等式和不等式) 端点约束函数可能不是局部 Lipschitz 的, 对于描述目标和不等式约束的函数, 在参考点附近甚至不是连续的, 然而, 只要问题中的函数在所考虑的点是 Fréchet可微的, 那么必要最优性条件以传统的 PMP 形式得到, 并且对特殊的非光滑函数类, 以上次微分扩张的形式得到.

与 6.1 节和 6.2 节中的逼近/扰动方法不同, 现在依赖更传统的本原空间方法, 该方法可追溯到对有限维控制系统 Pontryagin 最大值原理的经典证明 [124, 1102] 以及随后沿着 Rozonoér[1180] 所铺垫的道路上的重大发展. 所用的本原空间技巧有两个要素, 其踪迹可以在 McShane 的关于变分法的文章 [860] 中找到: 针形变分和凸分离. 这两个要素在最大值原理的证明 [124, 1102] 中是至关重要的, 然而, 在不同的方向上, 它们的澄清和重要修改始于 Rozonoér [1180] 以及 Dubovitskii 与 Milyutin [369,370]; 也请参见 1.4.1 小节和 6.5.1 小节中的其他参考文献和讨论.

在定理 6.37 中阐述的最大值原理的证明中, 主要遵循始于 Rozonoér[1180] 文章的路线, 该文章分为三部分. Rozonoér 可能是第一个完全意识到, 在最大值原理中, 自由端点"终端控制"(即 Mayer) 问题的关键的变分作用, 并且利用针形变分对这类问题的 PMP 的证明发展了增量方法. 于是, 如有限维空间中非线性数学规划那样, 端点约束由凸分离技巧来处理, 这关联到所谓的像空间分析; 请比照 Plotnikov [1083], Gabasov 与 Kirillova [485] 以及近期的书 Giannessi [504]. 由仅为可微的函数给出的等式端点约束控制问题横截性条件的微妙推导由 Halkin[545] 发展而得, 它是基于 Brouwer 不动点定理.

定理 6.38 中得到的 PMP 的上次可微条件似乎是新的, 即使对有限维控制系统亦如此. 最接近的条件在最近的著作 Gannarsa 与 Sinestrari [217, 定理 7.3.1] 中对有限维自由端点控制问题在更强的条件下导出; 相关结果由 Mordukhovich 和 Shvartsman[955, 956] 对离散时间和离散逼近导出; 参见 6.4 节. 注意到, 值函数的 Fréchet上次梯度 (或"超梯度") 在最优控制中的通过 Hamilton-Jacobi 方程的合成问题上用到; 例如参见 Subbotina [1231], Zhou [1366], Cannarsa 与 Frankowska [216], Cannarsa 与 Sinestrari [217], Frankowska [472] 及其参考文献.

6.5.22 像空间中的多针形变分和凸分离

在 6.3.2~6.3.4 小节中给出的定理 6.37 的证明主要发展了由 Gabasov 与 Kirillova [485] 中实施的方法, 但原文是针对有限维控制系统的, 并且假设更强. 如前所述, 6.3.2 小节中自由端点问题的基本证明思想源于 Rozonoér [1180], 而利用增

量公式的可测控制的针形变分则如 Mordukhovich [887, 901] 那样处理. 关于针形变分, 包括它们在高阶必要最优性条件中的应用, 读者可在 Agrachev 与 Sachkov [2], Bianchini 与 Kawski [114], Krener [703], Ledzewicz 与 Schättler[756], Sussmann [1236,1238] 以及其中的参考文献中找到更多最近的进展.

与自由端点问题相比, 定理 6.37 在有端点约束时的证明要复杂得多, 要求考虑到动态控制系统可达集的几何性质. 在约束的框架之下, 多针形变分的应用是至关重要的, 它使得可以在像空间中构造可达集的凸切向逼近, 其维数等于端点约束的数目加上费用函数的数目一. 于是, 虽然所考虑的控制问题涉及无穷维动态/状态空间, 但是最大值原理的证明依赖于有限维的凸分离.

注意到, 为了得到所要的点基结果, 6.3 节不涉及 SNC 类型的性质, 这不同于 6.1 节和 6.2 节中的一般情形. 实际上这与前面章节中得到的结果是一致的, 因为可以注意到, 约束/目标集的 SNC 性质在有限多个端点约束的情形自动成立. 该现象涉及约束集的有限余维数性质, 由此很容易就可导出在无穷维空间中不可避免的序列法紧性. 也注意到, 从 6.3.3 小节和 6.3.4 小节中的证明可以看到, 在像空间中相关逼近集的凸性源于时间区间的连续性; 这是连续时间控制系统所固有的隐藏凸性的另外一种表现.

6.5.23 离散最大值原理

6.4 节再次考虑具有离散时间和连续时间系统的离散逼近的最优控制问题, 但目的和 6.1 节是截然不同的. 在 6.1 节中, 虽然得到的离散包含的结果有其独立的意义, 但离散逼近主要是用来作为导出微分包含必要最优性条件的驱动力. 在 6.1.4 小节中, 对一般 (非凸和非 Lipschitz) 离散包含, 通过将其约化为具有许多几何约束的非凸数学规划问题, 建立了 Euler-Lagrange 类型的必要最优性条件. 当“离散速率”集 $F_j(x)$ 为凸时, 由定理 1.34 中关于凸值映射上导数的极值性质 (这实际上是由于凸集法锥的极值形式), 所得的结果自动地蕴涵最大值类型条件. 从非光滑分析的一般角度来看, 毋庸质疑, 某种特定的凸性对于这类极值类型的表示是必须的. 另一方面, Pontryagin 最大值原理及其非光滑推广对具有无显式凸性假设的连续时间控制系统是成立的. 从 6.1~6.3 节的结果和讨论中可看到, 这是由于连续时间动态中很强的固有的隐藏凸性.

考虑具有固定步长离散系统的最优控制问题, 在没有一些凸性的假设下, 没有理由预期这类最大值型结果. 然而, 离散控制问题 Pontryagin 最大值原理的确切类比首先由 Rozonoér[1180] 以离散最大值原理的名义得到, 该问题是在离散时间系统

$$\begin{cases} x(t+1) = Ax(t) + b(u(t), t), & x(0) = x_0, \\ u(t) \in U, & t = 0, \cdots, K-1 \end{cases} \tag{6.130}$$

上, 关于右端点 $x(K)$ 无任何约束并且无凸性假设, 极小化关于右端点 $x(K)$ 的线性函数. 这个结果的证明是基于在点 $t=\theta$ 的参考最优控制针形变分上的增量公式, 这类似连续时间情形, 但这当然不涉及 (不存在的) 区间 "小长度". 该结果及其证明由 Rozonoér 给出, 它在很大程度上依赖系统 (6.130) 的具体结构. 这同时也可能造成一个错觉, 即离散最大值原理可能对一般非线性系统, 至少是对自由端点问题成立. 注意到, 文献 [1180] 中明确表示了对这样的可能性的怀疑.

但有数篇文章 (大多是西方的文献) 以及在 Fan 和 Wang 的书 [426] 中, 错误地 "证明" 了离散最大值原理对最优性是必要的. 关于违反离散最大值原理的第一个明确 (相当复杂) 的例子由 Butkovsky [208] 给出. 在该方向上的许多其他例子可在 Gabasov 与 Kirillova [486] 的著作中找到, 它们比文献 [208] 中的要简单得多; 也可参见其中的参考文献.

例 6.46 取自 Mordukhovich [901]. 注意到, 它描述了一类离散控制系统, 其中整体极小值 (而不是极大值) 条件在初始条件之间的某种关系下成立. 文献 [901] 另外的例子表明了离散最大值原理可能不成立, 即使是对 (6.130) 类型的系统亦如此. 该系统关于状态和控制变量都是线性的, 并且具有非线性极小函数和非凸控制集 U. 这样一来就得到 Gabasov 与 Kirillova[486, 第5章的评论](后来由几位作者重复) 关于离散时间系统中具有充分小步长的离散最大值原理的正确性和连续时间系统最优解存在性之间的关系的猜想的反例. 在这个方向上更为引人注目的反例在 6.4.3 小节和 6.4.4 小节中给出, 它们表明了连续时间系统的最优控制的存在甚至并不蕴涵着离散逼近的最大值原理的近似版本.

关于非线性控制系统

$$\begin{cases} x(t+1)=f(x(t),u(t),t), & x(0)=x_0, \\ u(t)\in U, & t=0,\cdots,K-1, \end{cases} \tag{6.131}$$

离散最大值原理正确性的第一个正确结果可能要归功于 Halkin[540], 这是在容许 "速率集" $f(x,U,t)$ 的凸性假设下建立的; 对于该方向上更进一步的结果和讨论, 也请参见 Cannon, Cullum 与 Polak [218], Boltyanskii [127] 以及 Propoi [1105] 等书. 另一方面, Gabasov 与 Kirillova [486] 以及 Mordukhovich [901] 考虑了特殊类型的非线性自由端点控制问题, 在没有凸性的假设下证明了离散最大值原理. 更进一步, Mordukhovich 在书 [901] 中包括了针对离散最大值原理的所谓 "个别条件", 使得可以描述非凸系统初始数据之间的关系, 以确保离散最大值原理的正确或不正确性. 特别地, 这些条件可以完整地处理例 6.46 中的情形: 在那里, 离散最大值原理成立当且仅当 $\gamma\leqslant 0$ 和 $\eta\geqslant 0$.

6.5.24 自由端点离散参数系统的必要条件

上述讨论明显地表明了, 在最优控制的经典框架下, 连续时间及其离散时间

Pontryagin 最大值原理之间的差距, 即使对自由端点问题亦如此. 除了该现象的引人注目的理论价值外, 它还可能对数值分析有严重的影响, 因为它标志着 PMP 在计算之下可能的不稳定性, 而数值分析不可避免地要求对时间的离散化. 然而, 计算机的计算处理的不是固定步长的离散系统类型 (6.131), 而是下述参数类型的离散逼近系统

$$x(t+h) = x(t) + hf\big(x(t), u(t), t\big) \quad (\text{当 } h \downarrow 0), \tag{6.132}$$

其中的步长 h 为离散化参数. 因此, 自然要考虑涉及参数的系统 (6.132) 的控制问题的必要最优性条件, 它们自身依赖于参数 h.

该方向上的第一个结果由 Gabasov 与 Kirillova[484, 486] 得到, 他们在“拟最大值原理”的名称下, 在相当标准的光滑性但没有凸性的假设下, 导出了由一般离散时间系统

$$x(t+1) = f\big(x(t), u(t), t, h\big), \quad x \in \mathbb{R}^n, \quad h \in \mathbb{R}^m$$

控制的自由端点参数控制问题的必要最优性条件. 他们的结果断言, 对任何给定的 $\varepsilon > 0$, 在部分控制区域 (依赖于 ε 和 h) 上满足的某种 ε- 最大值条件. 具体到离散逼近系统 (6.132), 当 ε 和 h 越小时, ε- 最大值条件就越接近 Pontryagin 最大值条件. 随后, 在 Moiseev[884, 885] 以及 Ermoliev, Gulenko 与 Tzarenko[407] 的著作中, 对非凸自由端点控制问题的离散逼近导出了类似的结果; 在连续时间和离散时间系统的最优控制中, 对于这类结果在数值方法上的各种讨论和应用, 请参见上述著作以及 Propi [1105] 和 Evtushenko [412] 等书.

对于文献 [407, 484, 486, 884, 885] 中给出的离散逼近自由端点问题的拟最大值原理和相关结果的证明, 相互之间是类似的, 实际上, 都类似于关于轨道无约束的连续时间系统 PMP 的 Rozonoér的证明; 例如比较定理 6.37 在 6.3.2 小节中无约束情形的证明和 6.3.2 小节中光滑约束情形的定理 6.50 的证明. 所有这些证明很大程度上利用到控制问题的无约束性质, 并涉及关于最优控制单针形变分的费用增量公式. 连续时间和有限差分情形的唯一不同是, 在离散时间系统的参数族中利用的是小的离散化步长而不是离散时间系统中针形变分的小长度. 这些证明没有给得到具有端点约束最优控制问题的离散逼近相应 PMP 的可能性提供任何线索. 为建立这种 PMP, 隐藏凸性和可达集几何性质的一些相应的有限差分版本必须起到关键的作用.

6.5.25 约束离散逼近的近似最大值原理

具有光滑动态和光滑端点约束离散逼近 (6.132) 的最优控制问题的近似最大值原理 (AMP) 形式的必要最优性条件最初由 Mordukhovich 在文章 [891] 中提出, 接着在随后发表的文章 [899, 900, 901, 903, 942] 中进一步发展. 定理 6.59 给出的关于

光滑控制问题的最终形式是在文献 [901, 903] 中建立的; 也请参见文献 [906]. 这个主要定理在 6.4.5 小节中给出的证明是沿着本原空间的方向, 然而, 它在关键方面与 6.3.3 小节和 6.3.4 小节中考虑的连续时间版本有着显著的差异. 定理 6.59 中的 AMP 的证明有三个关键的假设:

— 等式约束扰动的一致性;

— 最优控制序列的恰当性;

— 关于状态变量初始数据的光滑性.

这些假设中的每一个对非凸约束问题离散逼近 AMP 的正确性是至关重要的, 这在 6.4.4 小节中由反例表明.

无论是从值收敛还是从 AMP 正确性的角度来看, 为实现离散逼近的稳定性, 端点约束一致扰动的作用是很关键的, 这由 Mordukhovich 从他一开始的研究就意识到了; 参见文献 [890, 891]. 例 6.61 表明, 如果端点约束没有被适当扰动 (必须比离散化步长递减得慢), 那么 AMP 可能不成立. 该例子取自 Mordukhovich 与 Raketskii [942]; 也参见文献 [901,903].

对约束非凸问题中沿着最优控制参考序列的 AMP 的正确性, 例 6.60 表明了恰当性性质的重要性. 该例子取自 Mordukhovich 与 Shvartsman [956]. 虽然该恰当性性质可视为对连续时间系统没有任何限制的分段连续性或可测控制的 Lebesgue 点的一些类似版本, 但它却是离散逼近所特有的. 注意到, 在自由端点问题中, 不需要作恰当性假设来确保 AMP; 参见定理 6.50 及其证明.

6.5.26 近似最大值原理的非光滑形式

近似最大值原理最为引人注目的特点之一是其对非光滑性的灵敏性. 这可能是关于最优性条件及变分分析相关议题的唯一没有任何常规下次微分 (关于极小化) 非光滑 (甚至凸) 情形推广的结果. 这由 Mordukhovich 与 Shvartsman 文章 [956] 中的例子表明, 这些例子在 6.4.3 小节对自由端点控制问题给出.

另一方面, 上述文章 [956] 证明了涉及上次微分横截性条件, 具有非凸费用函数自由端点问题 (定理 6.50) 和由非光滑函数描述的不等式类型端点约束的约束问题 (定理 6.66) 的近似最大值原理的新形式. 在这个方向上得到的结果适用于文献 [956] 中称为一致上次可微的特殊类非光滑函数. 除了光滑和凹函数外, 这类函数也包括半凹函数 (参见 5.5.4 小节), 实际上, 在 Nurmibskii [1017] 的意义下, 它与"弱凹"函数的局部化形式紧密相连, Nurmibskii 在数值最优化中有效地利用到它们. 定理 6.49 在自反空间中似乎是新的结果; 在有限维中, 它的一些结论和相关性质在文献 [956, 1017] 中建立, 但证明不同.

关于自由端点问题 AMP 的定理 6.50 给出了来自 Mordukhovich 与 Shvartsman [956] 的上次微分结果的无穷维推广, 它的光滑形式 [901] 实际上等价于 Gabasov 与

Kirillova[484, 486] 在较强假设下建立的"拟最大值原理".

注意到, 在光滑和非光滑情形, 定理 6.50 中自由端点形式的 AMP 都并不完全来自 6.4.4 小节中的约束形式. 对于自由端点问题, 除了无穷维数和没有恰当性性质外, 推论 6.52 和推论 6.53 中对最大值条件 (6.85) 有速度为 $\varepsilon(t, h_N) = O(h_N)$ 的误差估计, 这在任意 Banach 空间中对光滑和凹费用函数都成立.

6.5.27 近似最大值原理的应用

6.4.5 小节末给出了近似最大值原理的两个典型的应用. 在注 6.67 中描述的第一个, 遵循 Gabasov, Kirillova 与 Mordukhovich 的文章 [488], 利用离散逼近的值收敛和必要最优性条件导出连续时间系统的次最优性条件.

其次考虑了一个更为实际的应用, 利用近似最大值原理来解决由具有充分小步长离散时间系统的最优控制问题. 例 6.68 取自 Mordukhovich [901], 考虑的是一个(简化的) 化学工程实际问题, 该问题的描述见 Fan 与 Wang 的书 [426]. 离散最大值原理不能用来寻找这类约束非凸问题的最优解, 文献 [426] 的作者在其整个著作和相关文章中错误地这样做了. 另一方面, 定理 6.59 中近似最大值原理却可以用来找到最优控制.

约束离散逼近问题 AMP 的其他应用由 Nitka-Styczen[1013–1015] 给出, 她考虑的是涉及等式端点约束的最优周期控制. 根据 AMP 的机理, 她设计了解决这类问题的有效数值方法并将它们应用到来自化学、生物技术和生态过程最优化等实际问题中. 文献 [1015] 中考虑的一些模型是由遗传/时滞控制系统描述的, 要求对文献 [1015] 以及本书 6.4.6 小节给出的 AMP 阐述作某些修改.

6.5.28 时滞系统中的近似最大值原理

6.4.6 小节中给出的结果, 及其在无穷维空间中时滞系统的直接推广, 取自 Mordukhovich 与 Shvartsman [956]. 为了简单起见, 该小节只考虑自由端点问题, 对具有时滞状态变量的非线性系统, 导出了具有上次可微横截性条件的 AMP. 对于时滞系统, 遵循 Warga [1315] 的方法, 该结果的证明基于将时滞系统约化为普通的离散系统并且相关时间区间长度 $b - a$ 和离散化步长 h_N 之间可能不可公度.

6.4.6 小节最后的例 6.70 将读者的注意力吸引到一类非常有趣的遗传系统上, 称为中立型泛函–微分系统, 这与仅关于状态具有时滞的普通控制系统及其推广明显不同. 这类关于速率变量有时滞的系统会在 7.1 节中更详细地考虑; 也请参见第 7 章的评论. 例 6.70 是 Gabasov 与 Kirillova 的书 [485, 3.66节] 中连续时间例子的有限差分形式, 它表明了对于由中立型有限差分系统控制的光滑自由端点控制问题没有自然类似的 AMP.

第 7 章　分布系统的最优控制

本章继续从变分分析与广义微分高等方法的角度研究最优控制问题. 上一章致力于由常微分方程和包含控制的系统及其离散版本, 这里的主要精力放在由泛函微分和偏微分关系控制的分布参数系统. 特别地, 本章研究时滞微分–代数包含的最优控制, 这包括很多本质上不同于常微分控制的重要系统; 还讨论由双曲和抛物偏微分方程控制的系统, 这涉及 Dirichlet 和 Neumann 类型的边界控制以及逐点状态约束. 这里提到的问题在以前的文献中尚未有充分的研究. 本章大部分内容都基于作者及合作者近期的结果.

本章从所谓的时滞微分–代数系统的最优控制问题开始, 它描述的控制系统由互相关联的时滞微分包含和代数方程给出, 同时含有连续和离散时间系统的性质. 这特别包括第 6 章简要讨论的中立型的泛函微分控制系统. 接下来考虑具有逐点状态约束双曲系统的边界控制问题. 这从本质上比分布控制的系统要难一些 (因为没有正则性). 根据边界条件 (Neumann 和 Dirichlet 类型) 的不同, 这些问题也互不相同. 本章的最后一节研究不确定条件下抛物系统的极小极大控制问题, 它具有 Dirichlet 边界控制和逐点状态约束. 主要结果是建立变分分析框架下的必要最优和次最优条件, 以及一些逼近技术的相关收敛性/稳定性结果.

7.1　时滞微分–代数包含的优化

本节处理微分–代数控制系统的动态优化问题, 其所在的领域非常重要但尚有待深入探索. 从数学上看, 微分–代数系统提供了由互相关联的微分与代数的关系给出的控制过程的描述. 这种动态模型有很多应用, 特别包括过程系统工程、机器人、带完全和不完全约束的力学系统等; 见本章评注中的文献. 虽然由微分–代数控制的动态系统非常重要, 变分分析中对此类最优控制的研究工作尚很少, 特别是在必要最优和次最优条件上. 以前最好的工作是关于由微分–代数“方程”描述的控制过程的, 需要相当苛刻的“指标为 1”条件, 这在许多重要的应用中并不成立. 微分–代数系统的一个有趣的特性是, 在没有凸性假设时, 这种系统的最优过程不满足 Pontryagin 最大值原理自然的类似版本, 即使在指标为 1 时亦如此.

本节研究的微分–代数系统涉及微分包含关系, 这与以前研究中的方程是不同的. 另一方面, 这里的代数方程假设是线性的, 而不是要求指标为一. 一个主要的改进是在微分和代数关系中都引入了时滞, 这其实是一种正则化的因子, 使得可以在

代数方程中区分指标 1 和更高的项. 要研究的主要问题记作 (DA), 定义如下:

$$
\begin{cases}
\min \quad J[x,z] := \varphi\big(x(a),x(b)\big) + \displaystyle\int_a^b \vartheta\big(x(t),x(t-\theta),z(t),\dot z(t),t\big)\,\mathrm{d}t \\
\text{s.t.} \\
\qquad \dot z(t) \in F\big(x(t),x(t-\theta),z(t),t\big), \quad \text{a.e. } t\in[a,b], \\
\qquad z(t) = x(t) + Ax(t-\theta), \quad t\in[a,b], \\
\qquad x(t) = c(t), \quad t\in[a-\theta,a), \\
\qquad \big(x(a),x(b)\big) \in \Omega,
\end{cases}
$$

这里 $x\colon [a-\theta,b]\to X$ 在 $[a-\theta,a)$ 和 $[a,b]$ 上是连续的 (仅可能在 $t=a$ 处有一个跳跃), $z(\cdot)$ 在 $[a,b]$ 上是绝对连续的. 简单起见, 本节假设 $X=\mathbb{R}^n$, 即状态空间是“有限维”的. 根据 6.1 节中建立的方法, 可以把以下得到的结果扩展到无限维状态空间 X, 其中的假设可适当地根据 6.1 节中常微分发展包含关系的要求平行地给出. 需要指出, 即便在 $X=\mathbb{R}^n$ 的情形, (DA) 仍然是泛函微分控制系统的无限维优化问题, 这与常微分情形的对应问题是非常不同的.

当 F 不依赖于 z 时, (DA) 中的动态系统简化为“中立型”的泛函微分系统, 写成所谓的“Hale 类型”时即为

$$
\frac{\mathrm{d}}{\mathrm{d}t}\big[x(t)+Ax(t-\theta)\big] \in F\big(x(t),x(t-\theta),t\big), \quad \text{a.e. } t\in[a,b].
$$

因此上面描述的 Bolza 问题 (DA) 可以看成中立型系统最优化问题的拓展, 它对应于被积函数 ϑ 不依赖于 $(z,\dot z)$ 的情形. 要强调的是, 中立型系统 (以及此处考虑的更一般的微分–代数系统) 的动态优化问题比 $A=0$ 时的常微分和时滞微分系统情形在本质上要难得多, 表现出新的特点; 见后面的讨论.

此后总假设 $F\colon \mathbb{R}^n\times\mathbb{R}^n\times\mathbb{R}^n\times[a,b]\rightrightarrows\mathbb{R}^n$ 为具有闭图的集值映射, Ω 为闭集, $\theta>0$ 为正常数时滞, A 为 $n\times n$ 常数矩阵. 注意本节用的方法也可以来讨论“多时滞”$\theta_1\geqslant\theta_2\geqslant\cdots\geqslant\theta_m>0$, 以及“可变时滞”$\theta(\cdot)$ 的情形, 这里 $|\dot\theta(t)|<\alpha\in(0,1)$ 对 $t\in[a,b]$ 几乎处处成立.

正如 6.1 节中常微分包含的情形, 这里研究问题 (DA) 的方法基于“离散逼近方法”, 它在微分–代数包含的定性/数值计算/定量方面无疑都是很有意义的. 该方法在 (DA) 问题上的实施与 6.1 节在很多方面是不同的 (更复杂), 特别地, 它利用了非零的时滞 θ. 其中关键的问题是建立离散逼近的“变分稳定性”来保证最优解的恰当强收敛, 这和以前是一样的.

7.1.1 小节建立起 (DA) 中微分–代数动态系统的适定离散逼近, 没有考虑费用泛函和端点约束. 基本目标是用 (DA) 中的微分–代数包含有限差分版本的容许解来强逼近原包含的任意容许解 $\{x(\cdot),z(\cdot)\}$. 在 7.1.2 小节中, 应用这种强逼近, 就可

以对 (DA) 的离散逼近问题 (端点约束需要恰当的扰动) 的最优解作出相对于原问题的最优解的收敛性分析. 与常微分发展包含的情形一样, “松弛稳定性” 在所需的强变分收敛性的建立上至关重要.

应用变分分析中的广义微分工具, 7.1.3 小节导出了具有离散时间的差分–代数系统的必要最优条件, 该系统来自适定离散逼近. 状态空间的有穷维假设极大地简化了这个过程. 但是, 类似 6.1.4 小节中常微分发展系统的情形, 用前面建立的 SNC 分析以及对应的“模糊”结果还可以把这个过程推广到无穷维状态空间的情况. 最后, 通过对离散逼近取极限, 7.1.4 小节建立了微分–代数系统 (DA) 的以 Euler-Lagrange 及 Hamilton 包含的扩展形式给出的主要广义必要最优条件.

7.1.1 微分–代数包含的离散逼近

本小节处理时滞微分–代数系统 (DA) 任意容许解对的离散逼近, 此时不考虑费用泛函和端点约束. 在相当一般的假设下, 本节证明了, 用对应的以经典的 Euler 格式得到的有限差分包含系统的容许解对, 可以在下述意义下强逼近微分–代数系统的任意容许解对. 这个结果是构造性的, 它给出了逼近率的有效估计, 因此对时滞微分–代数系统的数值分析当然有独立的意义.

令 $\{\bar{x}(\cdot),\bar{z}(\cdot)\}$ 为 (DA) 中动态系统的一个“容许解对”, 即 $\bar{x}(\cdot)$ 在 $[a-\theta,a)$ 和 $[a,b]$ 上连续 (在 $t=a$ 点可能有一个跳跃), $\bar{z}(\cdot)$ 在 $[a,b]$ 上绝对连续, 并且满足 (DA) 中的动态关系. 注意, (DA) 中的端点约束对 $\{\bar{x}(\cdot),\bar{z}(\cdot)\}$ 不一定成立. 在成立的时候, 则这个解对是 (DA) 的“可行解”. 在整节中都利用如下的“常驻假设”:

(D1) 存在两个开集 $U\subset\mathbb{R}^n$, $V\subset\mathbb{R}^n$ 和两个正常数 ℓ_F, m_F 使得 $\bar{x}(t)\in U$ $(\forall t\in[a-\theta,b])$, $\bar{z}(t)\in V$ $(\forall t\in[a,b])$, $F(x,y,z,t)$ 是闭集, 且

$$\begin{aligned}&F(x,y,z,t)\subset m_F\mathbb{B},\quad \forall(x,y,z,t)\in U\times U\times V\times[a,b],\\&F(x_1,y_1,z_1,t)\subset F(x_2,y_2,z_2,t)\\&\qquad\qquad+\ell_F\big(\|x_1-x_2\|+\|y_1-y_2\|+\|z_1-z_2\|\big)\mathbb{B},\\&\qquad\forall(x_1,y_1,z_1),(x_2,y_2,z_2)\in U\times U\times V,\quad\forall t\in[a,b].\end{aligned}$$

(D2) $F(x,y,z,t)$ 对 $t\in[a,b]$ 几乎处处 Hausdorff 连续, 且对 $(x,y,z)\in U\times U\times V$ 是一致的.

(D3) 函数 $c(\cdot)$ 在 $[a-\theta,a]$ 上连续.

如 6.1.1 小节的 (6.6), 当 $(x,y,z)\in U\times U\times V$, 可对 $t\in[a,b]$ 上的 F 定义“连续性均模” $\tau(F;h)$, 如此根据命题 6.3 可知, (D2) 等价于 $\tau(F;h)\to 0$ (当 $h\to 0$).

把时滞微分–代数包含中的导数替换为“Euler 有限差分”, 即

$$\dot{z}(t)\approx\frac{z(t+h)-z(t)}{h},$$

下面构造该包含的一个离散逼近序列. 对任意 $N \in \mathbb{N}$, 考虑离散化步长 $h_N := \theta/N$, 并定义 $[a,b]$ 上的离散网

$$t_j := a + jh_N \quad (j = -N, \cdots, k), \quad t_{k+1} := b,$$

其中 k 是满足 $a + kh_N \leqslant b < a + (k+1)h_N$ 的自然数. 如此则 (DA) 中动态关系对应的“时滞差分–代数包含”可描述为

$$\begin{cases} z_N(t_{j+1}) \in z_N(t_j) + h_N F\big(x_N(t_j), x_N(t_j - \theta), z_N(t_j), t_j\big), \\ \qquad\qquad j = 0, \cdots, k, \\ z_N(t_j) = x_N(t_j) + Ax_N(t_j - \theta), \quad j = 0, \cdots, k+1, \\ x_N(t_j) = c(t_j), \quad j = -N, \cdots, -1. \end{cases} \tag{7.1}$$

给定满足 (7.1) 的二元组 $\{x_N(t_j), z_N(t_j)\}$, 考虑 $x_N(t_j)$ 在连续时间区间 $[a-\theta, b]$ 上的扩张, 使得 $x_N(t)$ 在 $[a,b]$ 上逐段线性, 在 $[a-\theta, a)$ 上逐段为常数且右连续. 再定义在 $[a,b]$ 上的逐段为常数的“离散速度扩张”

$$v_N(t) := \frac{z_N(t_{j+1}) - z_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}), \quad j = 0, \cdots, k.$$

记 $z_N(t) := x_N(t) + Ax_N(t-\theta)$, 则易知

$$z_N(t) = z_N(a) + \int_a^t v_N(r)\,\mathrm{d}r \quad t \in [a,b].$$

下面的结果是定理 6.4 对应于微分–代数系统的版本, 它保证了, 用满足 (7.1) 的二元组的扩张 $\{x_N(t), z_N(t)\}$ 可以强逼近 (DA) 中动态系统的任意容许解. 这里 $W^{1,2}[a,b]$ 是指 Sobolev 空间 $W^{1,2}\big([a,b];\mathbb{R}^n\big)$.

定理 7.1(微分–代数系统的强逼近)　设 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$ 为假设 (D1)~(D3) 下的 (DA) 中动态系统的容许解对. 则存在差分–代数系统 (7.1) 的解序列$\{\widehat{x}_N(t_j), \widehat{z}_N(t_j)\}$ 使得

$$\widehat{x}_N(t_0) = \bar{x}(a), \quad \forall N \in \mathbb{N},$$

且满足, 当 $N \to \infty$ 时, $\widehat{x}_N(t)$ $(a-\theta \leqslant t \leqslant b)$ 在 $[a-\theta, b]$ 上一致收敛到 $\bar{x}(\cdot)$, 同时 $\widehat{z}_N(t)$ $(a \leqslant t \leqslant b)$ 在 $[a,b]$ 上沿 $W^{1,2}$ 的强拓扑收敛到 $\bar{z}(t)$.

证明　由阶梯函数在 $L^1[a,b] := L^1\big([a,b];\mathbb{R}^n\big)$ 上的稠密性, 首先取序列 $\{w_N(\cdot)\}$ $(N \in \mathbb{N})$, 使得每个 $w_N(t)$ 在 $[t_j, t_{j+1})$ $(j = 0, \cdots, k)$ 上为常数, 且 $w_N(\cdot)$ 在 $N \to \infty$ 时沿 $L^1[a,b]$ 的范数拓扑收敛到 $\dot{\bar{z}}(\cdot)$. 类似定理 6.4 的证明可得 $w_N(\cdot)$ 的 $L^1[a,b]$-范数的估计 (6.7), 这就足以像 6.1.1 小节中常微分发展包含情形那样进行本定理的证明了. 为简化后面的计算, 设

$$\|w_N(t)\| \leqslant 1 + m_F, \quad \forall t \in [a,b], \quad N \in \mathbb{N}.$$

定义数项序列

$$\xi_N := \int_a^b \|w_N(t) - \dot{\bar{z}}(t)\| \,\mathrm{d}t \to 0 \quad (N \to \infty).$$

记 $w_{N_j} := w_N(t_j)$, 并递归定义 $\{u_N(t_j), s_N(t_j)\}$ 如下:

$$\begin{cases} u_N(t_j) := \bar{x}(t_j), & j = -N, \cdots, 0, \\ s_N(t_j) := u_N(t_j) + Au_N(t_j - \theta), & j = 0, \cdots, k+1, \\ s_N(t_{j+1}) := s_N(t_j) + h_N w_{N_j}, & j = 0, \cdots, k. \end{cases}$$

而扩张离散二元组 $\{u_N(t), s_N(t)\}$ 满足

$$\begin{cases} u_N(t) = \bar{x}(t_j), \quad \forall t \in [t_j, t_{j+1}), \quad j = -N, \cdots, -1, \\ s_N(t) = u_N(t) + Au_N(t - \theta), \quad \forall t \in [a, b], \\ s_N(t) = \bar{z}(a) + \displaystyle\int_a^t w_N(r) \,\mathrm{d}r, \quad \forall t \in [a, b]. \end{cases}$$

接下来证明 $u_N(\cdot)$ 在 $[a,b]$ 上一致收敛到 $\bar{x}(\cdot)$. 记 $y_N(t) := u_N(t) - \bar{x}(t)$, $\alpha_N(t) := \|y_N(t) + Ay_N(t-\theta)\|$. 对任意 $t \in [a,b]$, 有

$$\alpha_N(t) = \|s_N(t) - \bar{z}(t)\| \leqslant \int_a^t \|w_N(r) - \dot{\bar{z}}(r)\| \,\mathrm{d}r \leqslant \xi_N.$$

这蕴涵

$$\begin{aligned} \|y_N(t)\| &\leqslant \alpha_N(t) + \|A\| \cdot \|y_N(t-\theta)\| \\ &\leqslant \alpha_N(t) + \|A\|\alpha_N(t-\theta) + \|A\|^2 \cdot \|y_N(t-2\theta)\| \leqslant \cdots \\ &\leqslant \alpha_N(t) + \|A\|\alpha_N(t-\theta) + \cdots + \|A\|^m \alpha_N(t - m\theta) \\ &\quad + \|A\|^{m+1} \cdot \|y_N(t-(m+1)\theta)\|. \end{aligned}$$

由 (D3) 可看出 $c(\cdot)$ 在 $[a-\theta, a]$ 上一致连续. 任取序列 $\beta_N \downarrow 0$ $(N \to \infty)$, 则有

$$\|c(t') - c(t'')\| \leqslant \beta_N, \quad \forall t', t'' \in [t_j, t_{j+1}], \quad j = -N, \cdots, -1.$$

取满足 $a - \theta \leqslant b - (m+1)\theta < a$ 的整数 m, 则 $t - (m+1)\theta \in [t_j, t_{j+1})$ 对某个 $j \in \{-N, \cdots, -1\}$ 成立. 这蕴涵

$$\|y_N(t-(m+1)\theta)\| \leqslant \|c(t_j) - c(t-(m+1)\theta)\| \leqslant \beta_N.$$

因为 $m \in \mathbb{N}$ 不依赖于 N, 这给出

$$\|y_N(t)\| \leqslant \xi_N(1 + \|A\| + \cdots + \|A\|^m) + \|A\|^{m+1}\beta_N := \varrho_N \to 0 \quad (N \to \infty).$$

现在考虑由

$$\zeta_N := h_N \sum_{j=0}^{k} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)$$

定义的序列 $\{\zeta_N\}$ 并证明 $\zeta_N \downarrow 0\ (N \to \infty)$. 根据 ζ_N 和连续性均模 $\tau(F;h)$ 的构造, 可得如下估计:

$$\begin{aligned}
\zeta_N =& \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)\,\mathrm{d}t \\
=& \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t\big)\Big)\,\mathrm{d}t \\
&+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \Big[\operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big) \\
&- \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t\big)\Big)\Big]\,\mathrm{d}t \\
\leqslant& \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t\big)\Big)\,\mathrm{d}t + \tau(F; h_N).
\end{aligned}$$

进一步, 由 (D1) 可得, 对任意 $t \in [t_j, t_{j+1})$, $j = 0, \cdots, k$,

$$\begin{aligned}
&\operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t\big)\Big) \\
&- \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t), u_N(t-\theta), s_N(t), t\big)\Big) \\
\leqslant& \operatorname{dist}\Big(F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t\big); F\big(u_N(t), u_N(t-\theta), s_N(t), t\big)\Big) \\
\leqslant& \ell_F\Big(\big\|u_N(t_j) - u_N(t)\big\| + \big\|u_N(t_j-\theta) - u_N(t-\theta)\big\| + \big\|s_N(t_j) - s_N(t)\big\|\Big).
\end{aligned}$$

注意到

$$\begin{aligned}
\big\|s_N(t_j) - s_N(t)\big\| =& \Big\|\int_{t_j}^{t} w_N(r)\,\mathrm{d}r\Big\| \leqslant (1+m_F)(t_{j+1}-t_j) \\
=& (1+m_F)h_N := \eta_N \downarrow 0 \quad (N \to \infty),
\end{aligned}$$

则得估计

$$\begin{aligned}
&\big\|u_N(t) - u_N(t_j)\big\| \\
\leqslant& \eta_N + \|A\| \cdot \big\|u_N(t-\theta) - u_N(t_j-\theta)\big\| \\
\leqslant& \eta_N\big(1 + \|A\| + \cdots + \|A\|^m\big) + \|A\|^{m+1} \cdot \Big\|u_N\big(t-(m+1)\theta\big)
\end{aligned}$$

$$
\begin{aligned}
&-u_N(t_j-(m+1)\theta)\Big\|\\
\leqslant&\,\eta_N\big(1+\|A\|+\cdots+\|A\|^m\big)+\|A\|^{m+1}\beta_N:=\delta_N\downarrow 0\quad (N\to\infty).
\end{aligned}
$$

从而保证了

$$
\begin{aligned}
&\operatorname{dist}\Big(w_{N_j};F\big(u_N(t_j),u_N(t_j-\theta),s_N(t_j),t\big)\Big)\\
&-\operatorname{dist}\Big(w_{N_j};F\big(u_N(t),u_N(t-\theta),s_N(t),t\big)\Big)\\
\leqslant&(\eta_N+2\delta_N)\ell_F.
\end{aligned}
$$

由 (D1) 及上述估计可得, 对任意 $t\in[t_j,t_{j+1})$, $j=0,\cdots,k$, 有

$$
\begin{aligned}
&\operatorname{dist}\Big(w_{N_j};F\big(u_N(t),u_N(t-\theta),s_N(t),t\big)\Big)\\
&-\operatorname{dist}\Big(w_N(t);F\big(\bar{x}(t),\bar{x}(t-\theta),\bar{z}(t),t\big)\Big)\\
\leqslant&\operatorname{dist}\Big(F\big(u_N(t),u_N(t-\theta),s_N(t),t\big);F\big(\bar{x}(t),\bar{x}(t-\theta),\bar{z}(t),t\big)\Big)\\
\leqslant&\ell_F\Big(\big\|u_N(t)-\bar{x}(t)\big\|+\big\|u_N(t-\theta)-\bar{x}(t-\theta)\big\|+\big\|s_N(t)-\bar{z}(t)\big\|\Big)\\
\leqslant&(w\varrho_N+\xi_N)\ell_F.
\end{aligned}
$$

记 $\mu_N:=\eta_N+2\delta_N+2\varrho_N+\xi_N$, 则

$$
\begin{aligned}
&\operatorname{dist}\Big(w_{N_j};F\big(u_N(t_j),u_N(t_j-\theta),s_N(t_j),t\big)\Big)\\
\leqslant&\,\ell_F\mu_N+\operatorname{dist}\Big(w_{N_j};F\big(\bar{x}(t),\bar{x}(t-\theta),\bar{z}(t),t\big)\Big)\\
\leqslant&\,\ell_F\mu_N+\big\|w_{N_j}-\dot{\bar{z}}(t)\big\|.
\end{aligned}
$$

最后得

$$
\begin{aligned}
\zeta_N\leqslant&\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\Big(\big\|w_{N_j}-\dot{\bar{z}}(t)\big\|+\ell_F\mu_N\Big)\,\mathrm{d}t+\tau(F;h_N)\\
=&\,\xi_N+\ell_F\mu_N(b-a)+\tau(F;h_N):=\gamma_N\downarrow 0\quad(N\to\infty).
\end{aligned}\tag{7.2}
$$

注意离散二元组 $\{u_N(t_j),s_N(t_j)\}$ 未必是 (7.1) 的容许解. 利用“临近算法”, 构造

$$
\left\{
\begin{aligned}
&\widehat{x}_N(t_j)=c(t_j),\quad j=-N,\cdots,-1,\quad \widehat{x}_N(t_0)=\bar{x}(a),\\
&\widehat{z}_N(t_{j+1})=\widehat{z}_N(t_j)+h_Nv_{N_j},\quad j=0,\cdots,k,\\
&\widehat{z}_N(t_j)=\widehat{x}_N(t_j)+A\widehat{x}_N(t_j-\theta),\quad j=0,\cdots,k+1,\\
&v_{N_j}\in F\big(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\theta),\widehat{z}_N(t_j),t_j\big),\quad j=0,\cdots,k,\\
&\|v_{N_j}-w_{N_j}\|=\operatorname{dist}\Big(w_{N_j};F\big(\widehat{x}_N(t_j),\widehat{x}_N(t_j-\theta),\widehat{z}_N(t_j),t_j\big)\Big),\\
&\qquad\qquad j=0,\cdots,k.
\end{aligned}
\right.\tag{7.3}
$$

由此构造可以看出, $\{\widehat{x}_N(t_j), \widehat{z}_N(t_j)\}$ 对任意 $N \in \mathbb{N}$ 都是离散包含 (7.1) 的可行解. 注意到对 $j=-N,\cdots,-1$, $t \in [t_j, t_{j+1})$,

$$\|\widehat{x}_N(t)-\bar{x}(t)\| = \|\widehat{x}_N(t_j)-\bar{x}(t)\| = \|c(t_j)-c(t)\| < \beta_N.$$

这蕴涵 $\widehat{x}_N(\cdot)$ 的扩张在 $[a-\theta, a)$ 上一致收敛到 $\bar{x}(t)$.

现在分析 $[a,b]$ 上的情况. 首先应用数学归纳法证明 $\widehat{x}_N(t_j) \in U$, $\widehat{z}_N(t_j) \in V$ $(j=0,\cdots,k+1)$. 事实上, 显然有 $\widehat{x}_N(t_0) \in U$, $\widehat{z}_N(t_0) \in V$. 设 $\widehat{x}_N(t_j) \in U$, $\widehat{z}_N(t_j) \in V$ 对任意 $j=1,\cdots,m$ 成立, 这里 $m \in \{1,\cdots,k\}$ 是固定的. 则

$$\begin{aligned}
&\|\widehat{x}_N(t_{m+1})-u_N(t_{m+1})\| \\
=&\|\widehat{z}_N(t_{m+1})-A\widehat{x}_N(t_{m+1}-\theta)-s_N(t_{m+1})+Au_N(t_{m+1}-\theta)\| \\
\leqslant&\|A\|\cdot\|\widehat{x}_N(t_{m+1}-\theta)-u_N(t_{m+1}-\theta)\|+\|\widehat{z}_N(t_{m+1})-s_N(t_{m+1})\| \\
\leqslant&\|A\|\cdot\|\widehat{x}_N(t_{m+1}-\theta)-u_N(t_{m+1}-\theta)\| \\
&+\|A\|\cdot\|\widehat{x}_N(t_m-\theta)-u_N(t_m-\theta)\|+\|\widehat{x}_N(t_m)-u_N(t_m)\| \\
&+h_N\,\mathrm{dist}\Big(w_{N_m}; F\big(\widehat{x}_N(t_m),\widehat{x}_N(t_m-\theta),\widehat{z}_N(t_m),t_m\big)\Big).
\end{aligned}$$

考虑到估计

$$\begin{aligned}
&\|\widehat{x}_N(t_m)-u_N(t_m)\| \\
\leqslant&\|A\|\cdot\|\widehat{x}_N(t_{m-N})-u_N(t_{m-N})\| \\
&+\|A\|\cdot\|\widehat{x}_N(t_{m-1-N})-u_N(t_{m-1-N})\|+\|\widehat{x}_N(t_{m-1})-u_N(t_{m-1})\| \\
&+h_N\,\mathrm{dist}\Big(w_{N_{m-1}}; F\big(\widehat{x}_N(t_{m-1}),\widehat{x}_N(t_{m-1-N}),\widehat{z}_N(t_{m-1}),t_{m-1}\big)\Big), \\
&\mathrm{dist}\Big(w_{N_{m-1}}; F\big(\widehat{x}_N(t_{m-1}),\widehat{x}_N(t_{m-1-N}),\widehat{z}_N(t_{m-1}),t_{m-1}\big)\Big) \\
\leqslant&\mathrm{dist}\Big(w_{N_{m-1}}; F\big(u_N(t_{m-1}),u_N(t_{m-1-N}),s_N(t_{m-1}),t_{m-1}\big)\Big) \\
&+\ell_F\Big(\|\widehat{x}_N(t_{m-1})-u_N(t_{m-1})\|+\|\widehat{z}_N(t_{m-1})-s_N(t_{m-1})\| \\
&+\|\widehat{x}_N(t_{m-1-N})-u_N(t_{m-1-N})\|\Big), \\
&\|\widehat{z}_N(t_m)-s_N(t_m)\| \\
\leqslant&\|\widehat{x}_N(t_m)-u_N(t_m)\|+\|A\|\cdot\|\widehat{x}_N(t_{m-N})-u_N(t_{m-N})\|,
\end{aligned}$$

以及 $\|\widehat{x}_N(t_j)-u_N(t_j)\|=0$ $(j \leqslant 0)$, 则得

$$\begin{aligned}
&\|\widehat{x}_N(t_{m+1})-u_N(t_{m+1})\| \\
\leqslant&M_1h_N\sum_{j=0}^{m}\mathrm{dist}\Big(w_{N_j}; F\big(u_N(t_j),u_N(t_j-\theta),s_N(t_j),t_j\big)\Big) \\
\leqslant&M_1\gamma_N
\end{aligned} \tag{7.4}$$

对某个常数 $M_1 > 0$ 成立, 其中常数 γ_N 在 (7.2) 中定义 ($N \in \mathbb{N}$). 利用上述对 $\|y_N(t)\| = \|u_N(t) - \bar{x}(t)\|$ 的估计并在需要的时候增大 M_1, 就有

$$\big\|\widehat{x}_N(t_{m+1}) - \bar{x}(t_{m+1})\big\| \leqslant \varrho_N + M_1\gamma_N \to 0 \quad (N \to \infty).$$

这蕴涵 $\widehat{x}_N(t_j) \in U$ 对任意 $j = 0, \cdots, k+1$ 成立.

进一步可以看到

$$\begin{aligned}
&\big\|\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})\big\| \\
\leqslant &\big\|\widehat{z}_N(t_m) - s_N(t_m)\big\| + h_N\|v_{N_m} - w_{N_m}\| \\
\leqslant &\big\|\widehat{z}_N(t_m) - s_N(t_m)\big\| + h_N \operatorname{dist}\Big(w_{N_m}; F\big(\widehat{x}_N(t_m), \widehat{x}_N(t_m - \theta), \widehat{z}_N(t_m), t_m\big)\Big).
\end{aligned}$$

由此得

$$\begin{aligned}
&\big\|\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})\big\| \\
\leqslant &M_2 h_N \sum_{j=0}^{m} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j - \theta), s_N(t_j), t_j\big)\Big) \\
\leqslant &M_2\gamma_N
\end{aligned} \tag{7.5}$$

对某常数 $M_2 > 0$ 成立. 再注意到

$$\begin{aligned}
&\big\|\widehat{z}_N(t_{m+1}) - \bar{z}_N(t_{m+1})\big\| \\
\leqslant &\big\|\widehat{z}_N(t_{m+1}) - s_N(t_{m+1})\big\| + \big\|s_N(t_{m+1}) - \bar{z}(t_{m+1})\big\| \leqslant M_2\gamma_N + \xi_N.
\end{aligned}$$

这就保证了 $\widehat{z}_N(t_j) \in V$ 对任意 $j = 0, \cdots, k+1$ 成立.

余下需要证明序列 $\{\widehat{z}_N(\cdot)\}$ 在 $[a, b]$ 上沿 $W^{1,2}$ 的范数拓扑收敛到 $\bar{z}(\cdot)$, 即

$$\max_{t \in [a,b]} \big\|\widehat{z}_N(t) - \bar{z}(t)\big\| + \int_a^b \big\|\dot{\widehat{z}}_N(t) - \dot{\bar{z}}(t)\big\|^2 \mathrm{d}t \to 0 \quad (N \to \infty). \tag{7.6}$$

为此由 (7.4) 和 (7.5) 得到

$$\begin{aligned}
&\sum_{j=0}^{k+1} \big\|\widehat{x}_N(t_j) - u_N(t_j)\big\| \\
\leqslant &\sum_{j=0}^{k+1} M_1 \sum_{m=0}^{j-1} h_N \operatorname{dist}\Big(w_{N_m}; F\big(u_N(t_m), u_N(t_m - \theta), s_N(t_m), t_m\big)\Big) \\
\leqslant &M_1(b-a) \sum_{j=0}^{k} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j - \theta), s_N(t_j), t_j\big)\Big), \\
&\sum_{j=0}^{k+1} \big\|\widehat{z}_N(t_j) - s_N(t_j)\big\|
\end{aligned}$$

$$
\begin{aligned}
&\leqslant \sum_{j=0}^{k+1} M_2 \sum_{m=0}^{j-1} h_N \operatorname{dist}\Big(w_{N_m}; F\big(u_N(t_m), u_N(t_m-\theta), s_N(t_m), t_m\big)\Big)\\
&\leqslant M_2(b-a)\sum_{j=0}^{k} \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big).
\end{aligned}
$$

由 (D1) 和 (7.2)~(7.5), 这蕴涵

$$
\begin{aligned}
&\int_a^b \big\|\dot{\widehat{z}}_N(t) - w_N(t)\big\|\, dt = \sum_{j=0}^{k}\int_{t_j}^{t_{j+1}} \big\|\dot{\widehat{z}}_N(t) - w_N(t)\big\|\, dt\\
=&\sum_{j=0}^{k} h_N \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)\\
&+\sum_{j=0}^{k} h_N \Big[\operatorname{dist}\Big(w_{N_j}; F\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j-\theta), \widehat{z}_N(t_j), t_j\big)\Big)\\
&-\operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)\Big]\\
\leqslant&\sum_{j=0}^{k} h_N \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j), u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)\\
&+\sum_{j=0}^{k} \ell_F h_N \Big[\big\|\widehat{x}_N(t_j) - u_N(t_j)\big\| + \big\|\widehat{x}_N(t_j-\theta) - u_N(t_j-\theta)\big\|\\
&+\big\|\widehat{z}_N(t_j) - s_N(t_j)\big\|\Big]\\
\leqslant&\gamma_N + 2(M_1+M_2)(b-a)\ell_F \sum_{j=0}^{k} h_N \operatorname{dist}\Big(w_{N_j}; F\big(u_N(t_j-\theta), s_N(t_j), t_j\big)\Big)\\
\leqslant&\gamma_N + 2(M_1+M_2)\ell_F(b-a)\gamma_N.
\end{aligned}
$$

最后这个估计保证了

$$
\begin{aligned}
\int_a^b \big\|\dot{\widehat{z}}_N(t) - \dot{z}(t)\big\|\, \mathrm{d}t &\leqslant \int_a^b \big\|\dot{\widehat{z}}_N(t) - w_N(t)\big\|\, \mathrm{d}t + \int_a^b \big\|w_N(t) - \dot{z}(t)\big\|\, \mathrm{d}t\\
&\leqslant \gamma_N\big(1 + 2(M_1+M_2)(b-a)\ell_F\big) + \xi_N.
\end{aligned}
$$

结合 (D1) 和 (7.3), 这给出 $\|\dot{\widehat{z}}_N(t)\| \leqslant m_F$ 和 $\|\dot{z}(t)\| \leqslant m_F$. 从而

$$
\begin{aligned}
&\int_a^b \big\|\dot{\widehat{z}}_N(t) - \dot{z}(t)\big\|^2 \mathrm{d}t = \int_a^b \big\|\dot{\widehat{z}}_N(t) - \dot{z}(t)\big\| \cdot \big\|\dot{\widehat{z}}_N(t) + \dot{z}(t)\big\|\, \mathrm{d}t\\
\leqslant& 2m_F\Big[\gamma_N\big(1 + 2(M_1+M_2)(b-a)\ell_F\big) + \xi_N\Big] \downarrow 0 \quad (N \to \infty).
\end{aligned}
$$

最后注意到

$$\max_{t\in[a,b]} \left\|\widehat{z}_N(t)-\bar{z}(t)\right\|^2 \leqslant (b-a)\int_a^b \left\|\dot{\widehat{z}}_N(t)-\dot{\bar{z}}(t)\right\|^2 \mathrm{d}t,$$

即得 (7.6). 证毕. △

7.1.2 离散逼近的强收敛

本小节的目标是对动态优化问题 (DA) 构造一个适定离散逼近序列, 使得离散逼近问题的最优解在下述意义下强收敛到原来由时滞微分–代数包含控制的问题的一个给定最优解. 下述的构造类似 6.1.3 小节中常微分发展包含的情形, 它显式地涉及所考虑问题 (DA) 的最优解 $\{\bar{x}(\cdot),\bar{z}(\cdot)\}$, 接下来的小节将得到该问题的必要最优条件. 从证明可以看出, 所得结果对“松弛中间局部极小点”也成立 (比照 6.1.2 小节和 6.1.3 小节), 这里只限于全局解/绝对 (其实是强) 极小点以简化叙述.

对任意自然数 N, 考虑下面的“离散时间”动态优化问题 (DA_N):

$$\begin{aligned}\min\quad J_N[x_N,z_N] := {}& \varphi\big(x_N(t_0),x_N(t_{k+1})\big)+\|x_N(t_0)-\bar{x}(a)\|^2\\ &+h_N\sum_{j=0}^{k}\vartheta\Big(x_N(t_j),x_N(t_j-\theta),z_N(t_j),\frac{z_N(t_{j+1})-z_N(t_j)}{h_N},t_j\Big)\\ &+\sum_{j=0}^{k}\int_{t_j}^{t_{j+1}}\left\|\frac{z_N(t_{j+1})-z_N(t_j)}{h_N}-\dot{\bar{z}}(t)\right\|^2\mathrm{d}t.\end{aligned} \tag{7.7}$$

约束包括由时滞差分–代数包含 (7.1) 控制的“动态约束”“扰动端点约束”

$$\big(x_N(t_0),x_N(t_{k+1})\big)\in \Omega_N := \Omega+\eta_N\mathbb{B} \tag{7.8}$$

(其中 $\eta_N := \|\widehat{x}_N(t_{k+1})-\bar{x}(b)\|$, 而 $\widehat{x}_N(\cdot)$ 是来自定理 7.1 的 $\bar{x}(\cdot)$ 的逼近), 以及“辅助约束”(对某 $\varepsilon>0$)

$$\|x_N(t_j)-\bar{x}(t_j)\|\leqslant\varepsilon,\quad \|z_N(t_j)-\bar{z}(t_j)\|\leqslant\varepsilon,\quad j=1,\cdots,k+1. \tag{7.9}$$

最后这个辅助约束只用来保证 (DA_N) 最优解的存在, 推导必要最优条件时可以忽略; 见后述.

下述讨论中, 选取 (7.9) 中的 $\varepsilon>0$ 使得 $\bar{x}(t)+\varepsilon\mathbb{B}\subset U$ $(\forall t\in[a-\theta,b])$ 和 $\bar{z}(t)+\varepsilon\mathbb{B}\subset V$ $(\forall t\in[a,b])$, 选取足够大的 N 使得 $\eta_N<\varepsilon$. 注意问题 (DA_N) 具有“可行”解, 这是因为定理 7.1 中的二元组 $\{\widehat{x}_N(\cdot),\widehat{z}_N(\cdot)\}$ 满足 (7.1), (7.8), (7.9) 这些约束. 这样一来, 若在 (D1)~(D3) 的基础上加上下述假设 (D4):

(D4) φ 在 $U\times U$ 上连续, $\vartheta(x,y,z,v,\cdot)$ 在 $t\in[a,b]$ 上几乎处处连续, 且对 $(x,y,z,v)\in U\times U\times V\times m_F\mathbb{B}$ 一致, $\vartheta(\cdot,\cdot,\cdot,\cdot,t)$ 在 $U\times U\times V\times m_F\mathbb{B}$ 上连续, 且对 $t\in[a,b]$ 一致, Ω 在 $\big(\bar{x}(a),\bar{x}(b)\big)$ 附近局部闭.

则根据经典的 Weierstrass 定理, 每个 (DA_N) 都具有一个最优解对 $\{\bar{x}_N(\cdot), \bar{z}_N(\cdot)\}$.

下面证明 $\{\bar{x}_N(\cdot), \bar{z}_N(\cdot)\}$ 在定理 7.1 的意义下强收敛于 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$. 为此需要原问题 (DA) 的一个重要的内蕴性质 ——“松弛稳定性”(比照 6.1.2 小节). 在研究 (DA) 中的时滞微分–代数系统这个原问题的同时, 考虑其“凸化”问题:

$$\begin{aligned}&\dot{z}(t) \in \operatorname{co} F\big(x(t), x(t-\theta), z(t), t\big), \quad \text{a.e.} \quad t \in [a,b],\\&z(t) = x(t) + Ax(t-\theta), \quad t \in [a,b].\end{aligned} \tag{7.10}$$

进一步, 对 (DA) 中给定的被积函数 ϑ, 对任意 (x,y,z,t) 取其在集合 $F(x,y,z,t)$ 上的限制

$$\vartheta_F(x,y,z,v,t) := \vartheta(x,y,z,v,t) + \delta\big(v; F(x,y,z,t)\big).$$

记 ϑ_F 相对于变量 v 的凸化为 $\widehat{\vartheta}_F(x,y,z,v,t)$, 并对时滞微分–代数系统定义“松弛广义 Bolza 问题”($\overline{\mathrm{DA}}$) 如下:

$$\min \widehat{J}[x,z] := \varphi\big(x(a), x(b)\big) + \int_a^b \widehat{\vartheta}\big(x(t), x(t-\theta), z(t), \dot{z}(t), t\big)\,\mathrm{d}t, \tag{7.11}$$

其中的可行解对 $\{x(\cdot), z(\cdot)\}$ 具有与 (DA) 相同的尾部和端点约束. ($\overline{\mathrm{DA}}$) 的可行解对称为 (DA) 的一个“松弛解对”.

对松弛问题和原问题的费用函数的最优值, 显然有 $\inf(\overline{DA}) \leqslant \inf(DA)$. 如果

$$\inf(DA) = \inf(\overline{DA}),$$

则称原问题 (DA) 对松弛是稳定的. 当集合 $F(x,y,z,t)$ 和被积函数 ϑ 对变量 v 为凸时这显然成立. 但这个性质是远远超出凸的范围的; 比照 6.1.2 小节对常微分发展包含的讨论. 事实上, 从松弛稳定性这个角度看, 常微分系统和仅在状态变量有时滞的常微分系统是没有区别的. 但对中性和微分–代数系统却非如此. 建议读者参阅 Kisielewicz 的书 [682], 那里有讨论保证中性泛函–微分系统松弛稳定性的条件, 该系统具有非凸的速度集合. 类似的结果对所考虑的微分–代数系统也成立.

下面建立离散逼近问题最优解的强收敛定理, 它是连接时滞微分–代数和差分–代数系统控制的最优控制问题的桥梁.

定理 7.2(差分–代数逼近最优解的强收敛)　设 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$ 为问题 (DA) 的一个最优解对, 并设其对松弛是稳定的, 假设 (D1)~D(4) 均成立. 则 (DA_N) 的任意最优解对序列 $\{\bar{x}_N(\cdot), \bar{z}_N(\cdot)\}$ ($N \in \mathbb{N}$) 都可以分别扩张到 $[a-\theta, b]$ 和 $[a,b]$ 上, 并在 $N \to \infty$ 时强收敛到 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$, 即 $\bar{x}_N(\cdot)$ 在 $[a-\theta, b]$ 上一致收敛到 $\bar{x}(\cdot)$, $\bar{z}_N(\cdot)$ 在 $[a,b]$ 上沿 $W^{1,2}$ 的范数拓扑收敛到 $\bar{z}(\cdot)$.

证明　由上述讨论, (DA_N) 在 N 足够大时总有最优解 $\{\bar{x}_N(\cdot), \bar{z}_N(\cdot)\}$, 因此不失一般性, 可设对任意 $N \in \mathbb{N}$ 这都成立. 对给定的最优解 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$ 应用定理 7.1

中的强逼近结果, 可得序列 $\{\widehat{x}_N(\cdot), \widehat{z}_N(\cdot)\}$. 因为任何 $\{\widehat{x}_N(\cdot), \widehat{z}_N(\cdot)\}$ 都是 (DA_N) 的可行解, 则有

$$J_N\left[\bar{x}_N, \bar{z}_N\right] \leqslant J_N\left[\widehat{x}_N, \widehat{z}_N\right], \quad \forall N \in \mathbb{N}.$$

方便起见, 把 $J_N\left[\widehat{x}_N, \widehat{z}_N\right]$ 写成三项的和:

$$\begin{aligned} J_N\left[\widehat{x}_N, \widehat{z}_N\right] = I_1 + I_2 + I_3 := & \varphi\big(\widehat{x}_N(t_0), \widehat{x}_N(t_{k+1})\big) \\ & + h_N \sum_{j=0}^{k} \vartheta\Big(\widehat{x}_N(t_j), \widehat{x}_N(t_j - \theta), \widehat{z}_N(t_j), \frac{\widehat{z}_N(t_{j+1}) - \widehat{z}_N(t_j)}{h_N}, t_j\Big) \\ & + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\| \frac{\widehat{z}_N(t_{j+1}) - \widehat{z}_N(t_j)}{h_N} - \dot{z}(t) \right\|^2 \mathrm{d}t. \end{aligned}$$

由定理 7.1 和 (D4) 中对 φ 的假设得

$$I_1 \to \varphi\big(\bar{x}(a), \bar{x}(b)\big) \quad (N \to \infty).$$

下面的目标是证明

$$\limsup_{N \to \infty} J_N\left[\bar{x}_N, \bar{z}_N\right] \leqslant J[\bar{x}, \bar{z}]. \tag{7.12}$$

这显然蕴涵于极限关系

$$J_N\left[\widehat{x}_N, \widehat{z}_N\right] \to J[\bar{x}, \bar{z}] \quad (N \to \infty).$$

为证明此式, 需要计算上述 $J_N\left[\widehat{x}_N, \widehat{z}_N\right]$ 表述中 I_2, I_3 的极限. 以 $\sim$ 表示在 $N \to \infty$ 时表达式的等价性, 并记

$$\widehat{v}_N(t) := \frac{\widehat{z}_N(t_{j+1}) - \widehat{z}_N(t_j)}{h_N}, \quad t \in [t_j, t_{j+1}), \quad j = 0, \cdots, k,$$

则有关系

$$\begin{aligned} I_2 = & h_N \sum_{j=0}^{k} \vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j - \theta), \widehat{z}_N(t_j), \widehat{v}_N(t_j), t_j\big) \\ = & \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j - \theta), \widehat{z}_N(t_j), \widehat{v}_N(t), t\big)\, \mathrm{d}t \\ & + \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \Big[\vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j - \theta), \widehat{z}_N(t_j), \widehat{v}_N(t), t_j\big) \\ & - \vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j - \theta), \widehat{z}_N(t_j), \widehat{v}_N(t), t\big)\Big]\, \mathrm{d}t \end{aligned}$$

$$
\begin{aligned}
&= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j-\theta), \widehat{z}_N(t_j), \widehat{v}_N(t), t\big)\,\mathrm{d}t + \tau(\vartheta; h_N) \\
&\sim \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \vartheta\big(\widehat{x}_N(t_j), \widehat{x}_N(t_j-\theta), \widehat{z}_N(t_j), \widehat{v}_N(t), t\big)\,\mathrm{d}t \\
&\to \int_a^b \vartheta\big(\bar{x}(t), \bar{x}(t-\theta), \bar{z}(t), \dot{\bar{z}}(t), t\big)\,\mathrm{d}t \quad (N\to\infty), \\
I_3 &= \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \big\|\widehat{v}_N(t) - \dot{\bar{z}}(t)\big\|^2\,\mathrm{d}t = \int_a^b \big\|\widehat{v}_N(t) - \dot{\bar{z}}(t)\big\|^2\,\mathrm{d}t \\
&= \int_a^b \big\|\dot{\widehat{z}}_N(t) - \dot{\bar{z}}(t)\big\|^2\,\mathrm{d}t \to 0 \quad (N\to\infty).
\end{aligned}
$$

这就蕴涵所需不等式 (7.12).

进一步, 可以看到定理中的强收敛性可由下式得到

$$
\beta_N := \|\bar{x}_N(a) - \bar{x}(a)\|^2 + \int_a^b \|\dot{\bar{z}}_N(t) - \dot{\bar{z}}(t)\|^2\,\mathrm{d}t \to 0 \quad (N\to\infty).
$$

如果该式不成立, 则存在 $\beta > 0$ 和序列 $\{N_m\} \subset \mathbb{N}$ 使得 $\beta_{N_m} \to \beta$ $(m\to\infty)$. 基于 (7.1) 应用标准的紧性推理, 并利用在状态空间有穷维框架下 (D1) 中的有界假设, 找到一个绝对连续映射 $\widetilde{z}\colon [a,b] \to \mathbb{R}^n$, 和在 $[a-\theta, a)$ 及 $[a,b]$ 上连续的映射 $\widetilde{x}\colon [a-\theta, b] \to \mathbb{R}^n$, 使得 (不失一般性)$N\to\infty$ 时在 $L^2[a,b]$ 的弱拓扑下 $\dot{\bar{z}}_N(t) \to \dot{\widetilde{z}}(t)$, 对 $t \in [a-\theta, b]$ 一致有 $\bar{x}_N(t) \to \widetilde{x}(t)$, 并且 $\widetilde{z}(t) = \widetilde{x}(t) + A\widetilde{x}(t-\theta)$ $(\forall t \in [a,b])$. 根据经典的 Mazur 定理, 存在一个由 $\dot{\bar{z}}_N(t)$ 的凸组合构成的序列在 $L^2[a,b]$ 的范数拓扑下收敛于 $\dot{\widetilde{z}}(t)$. 进而就有一个子序列对 $t \in [a,b]$ 几乎处处逐点收敛到 $\dot{\widetilde{z}}(t)$. 因此

$$
\begin{cases}
\dot{\widetilde{z}}(t) \in \operatorname{co} F\big(\widetilde{x}(t), \widetilde{x}(t-\theta), \widetilde{z}(t), t\big), & \text{a.e. } t \in [a,b], \\
\widetilde{z}(t) = \widetilde{x}(t) + A\widetilde{x}(t-\theta), & t \in [a,b].
\end{cases}
$$

因为 $\widetilde{x}(\cdot)$ 显然满足 (DA) 中的初始尾部条件和端点约束, 所以它是松弛问题 $(\overline{\text{DA}})$ 的可行解. 根据假设就可以注意到

$$
\begin{aligned}
&h_N \sum_{j=0}^{k} \vartheta\Big(\bar{x}_N(t_j), \bar{x}_N(t_j-\theta), \bar{z}_N(t_j), \frac{\bar{z}_N(t_{j+1}) - \bar{z}_N(t_j)}{h_N}, t_j\Big) \\
&= \sum_{j=0}^{i} \int_{t_j}^{t_{j+1}} \vartheta\big(\bar{x}_N(t_j), \bar{x}_N(t_j-\theta), \bar{z}_N(t_j), \dot{\bar{z}}_N(t), t_j\big)\,\mathrm{d}t \\
&\to \int_a^b \vartheta\big(\widetilde{x}(t), \widetilde{x}(t-\theta), \widetilde{z}(t), \dot{\widetilde{z}}(t), t\big)\,\mathrm{d}t \quad (N\to\infty).
\end{aligned}
$$

还可以看到, 积分泛函

$$I[v] := \int_a^b \left\|v(t) - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t$$

在 $L^2[a,b]$ 的弱拓扑下是下半连续的, 这是因为其被积式对 v 是凸的. 由于

$$\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\|\frac{\bar{z}_N(t_{j+1}) - \bar{z}_N(t_j)}{h_N} - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t = \int_a^b \left\|\dot{\bar{z}}_N(t) - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t,$$

则有

$$\int_a^b \left\|\dot{\widetilde{z}}(t) - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t \leqslant \liminf_{N\to\infty} \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\|\frac{\bar{z}_N(t_{j+1}) - \bar{z}_N(t_j)}{h_N} - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t.$$

由此并以 (7.7) 中费用函数的形式对 $J_N[\bar{x}_N, \bar{z}_N]$ 在 $N \to \infty$ 时取极限, 则有不等式

$$J[\widetilde{x}, \widetilde{z}] + \beta \leqslant \lim_{N\to\infty} J_N[\bar{x}_N, \bar{z}_N].$$

从而由 (7.12) 得

$$J[\widetilde{x}, \widetilde{z}] \leqslant J[\bar{x}, \bar{z}] - \beta < J[\bar{x}, \bar{z}] \quad (若\beta > 0).$$

由松弛稳定性的假设, 这与解对 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$ 对于松弛问题 $(\overline{\mathrm{DA}})$ 的最优性矛盾. 所以 $\beta = 0$. 证毕. △

需要指出, 类似 6.1.3 小节, 当 F 关于 $t \in [a,b]$ 为可测时, 对定理 7.2 进行修改则也可以对微分–代数系统建立定理 6.14 中关于离散逼近的“值收敛”的类似结果.

7.1.3 差分–代数系统的必要最优条件

本小节推导离散逼近问题 (DA_N) 的必要最优条件, 这是通过将其化为具有多个几何约束的非光滑数学规划问题来实现的. 由于假设状态空间 $X = \mathbb{R}^n$ 是有穷维的, 这里不需要像 6.1.4 小节中用的 SNC 分析法则和/或模糊分析. 记

$$w := (x_0, \cdots, x_{k+1}, z_0, \cdots, z_{k+1}, v_0, \cdots, v_k) \in \mathbb{R}^{n(3k+5)},$$

并基于逼近问题 (DA_N)(进而也是原问题 (DA)) 的初始数据定义如下映射和集合:

$$\begin{aligned}\varphi_0(w) := &\varphi(x_0, x_{k+1}) + \|x_0 - \bar{x}(a)\|^2 + h_N \sum_{j=0}^{k} \vartheta(x_j, x_{j-N}, z_j, v_j, t_j)\\ &+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\|v_j - \dot{\bar{z}}(t)\right\|^2 \mathrm{d}t,\end{aligned}$$

$$\varphi_j(w) := \begin{cases} \|x_j - \bar{x}(t_j)\| - \varepsilon, & j = 1, \cdots, k+1, \\ \|z_{j-k-1} - \bar{z}(t_{j-k-1})\| - \varepsilon, & j = k+2, \cdots, 2k+2, \end{cases}$$

$$\Lambda_j := \big\{(x_0, \cdots, v_k) \mid v_j \in F(x_j, x_{j-N}, z_j, t_j)\big\}, \quad j = 0, \cdots, k,$$

$$\Lambda_{k+1} := \big\{(x_0, \cdots, v_k) \mid (x_0, x_{k+1}) \in \Omega_N\big\},$$

$$g_j(w) := z_{j+1} - z_j - h_N v_j, \quad j = 0, \cdots, k,$$

$$s_j(w) := z_j - x_j - A x_{j-N}, \quad j = 0, \cdots, k+1,$$

其中 $x_j := c(t_j)$ (当 $j < 0$). 则每个问题 (DA_N) 都等价地化为 $\mathbb{R}^{n(3k+5)}$ 上的非光滑数学规划问题 (MP) 如下, 它具有 $(k+2)$ (有限) 个几何约束:

$$\begin{cases} \min & \varphi_0(w) \\ \text{s.t.} & \varphi_j(w) \leqslant 0, \quad j = 1, \cdots, r, \\ & f(w) = 0, \\ & w \in \Lambda_j, \quad j = 0, \cdots, l, \end{cases}$$

其中 $r = 2k+2$, $l = k+1$, $f \colon \mathbb{R}^{n(3k+5)} \to \mathbb{R}^{2k+3}$ 定义为

$$f(w) := \big(g_0(w), \cdots, g_k(w), s_0(w), \cdots, s_{k+1}(w)\big), \quad w \in \mathbb{R}^{n(3k+5)}.$$

简化起见, 这里在符号中省略了 (DA_N) 的解和对应的对偶元对逼近序列下标 N 的依赖关系.

令 $\bar{w}$ 为对应于 (DA_N) 的问题 (MP) 的一个最优解. 在后面的讨论中假设 φ_0 和 $\vartheta(\cdot, t)$ 是局部 Lipschitz 连续的. 对 (MP) 应用命题 6.16 中的必要最优条件 (有穷维的情形), 并分开考虑 f 的分量 g_j 和 s_j 的 (向量) 乘子, 可得 $\mu_j \in \mathbb{R}$ $(j = 0, \cdots, 2k+2)$, $e_j^* \in \mathbb{R}^n$ $(j = 0, \cdots, k)$, $d_j^* \in \mathbb{R}^n$ $(j = 0, \cdots, k+1)$ 和 $w_j^* \in \mathbb{R}^{n(3k+5)}$ $(j = 0, \cdots, k+1)$ 满足

$$\begin{cases} \mu_j \geqslant 0, \quad j = 0, \cdots, 2k+2, \\ \mu_j \varphi_j(\bar{w}) = 0, \quad j = 1, \cdots, 2k+2, \\ w_j^* \in N(\bar{w}; \Lambda_j), \quad j = 0, \cdots, k+1, \\ -\displaystyle\sum_{j=0}^{k+1} w_j^* \in \partial\Big(\sum_{j=0}^{2k+2} \mu_j \varphi_j\Big)(\bar{w}) + \sum_{j=0}^{k} \nabla g_j(\bar{w})^* e_j^* + \sum_{j=0}^{k+1} \nabla s_j(\bar{w})^* d_j^*. \end{cases} \tag{7.13}$$

把 w_j^* 写为 $w_j^* = (x_{0j}^*, \cdots, x_{k+1\,j}^*, z_{0j}^*, \cdots, z_{k+1\,j}^*, v_{0j}^*, \cdots, v_{kj}^*)$. 可以看到, 每个 w_j^* 都恰好只有一个分量非零, 并且非零的分量满足

$$\big(x_{jj}^*, x_{j-N\,j}^*, z_{jj}^*, v_{jj}^*\big) \in N\big((\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j); \operatorname{gph} F(t_j)\big), \quad \forall j = 0, \cdots, k.$$

类似地可以看到, $w_{k+1}^* \in N\big(\bar{z}^N; \Lambda_{k+1}\big)$ 等价于

$$(x_{0\,k+1}^*, x_{k+1\,k+1}^*) \in N\big((\bar{x}_0, \bar{x}_{k+1}); \Omega_N\big),$$

而 w_{k+1}^* 的所有其他分量皆为零. 由 φ_j $(j = 1, \cdots, 2k+2)$ 的构造和定理 7.2 中离散最优解的强收敛性可得

$$\varphi_j(\bar{w}) < 0 \quad (j = 1, \cdots, 2k+2, \quad N \to \infty).$$

由 (7.13) 中的互补松弛性条件, 就有 $\mu_j = 0$ 对所有的这些下标都成立. 对剩下的下标 $(j = 0)$ 令 $\lambda := \mu_0$. 进一步可得, 对 (MP) 中的 g_j 和 s_j 有

$$\sum_{j=0}^{k} \nabla g_j(\bar{w})^* e_j^* = (0, \cdots, 0, e_0^*, e_0^* - e_1^*, e_{k-1}^* - e_k^*, e_k^*, -h_N e_0^*, \cdots, -h_N e_k^*),$$

$$\sum_{j=0}^{k+1} \nabla s_j(\bar{w})^* d_j^* = (-d_0^* + A^* d_N^*, -d_1^* + A^* d_{N+1}^*, \cdots, -d_{k-N+1}^* + A^* d_{k+1}^*,$$
$$-d_{k-N+2}^*, \cdots, -d_{k+1}^*, d_0^*, \cdots, d_{k+1}^*, 0, \cdots, 0).$$

把定理 2.33(c) 中的次微分和法则应用于 (MP) 中的"Lipschitz 和" φ_0, 则有

$$\partial\varphi_0(\bar{w}) \subset \partial\varphi(\bar{x}_0, \bar{x}_{k+1}) + 2\big(\bar{x}_0 - \bar{x}(a)\big) + h_N \sum_{j=0}^{k} \partial\vartheta(\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j, t_j)$$
$$+ \sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} 2\big(\bar{v}_j - \dot{\bar{z}}(t)\big)\,\mathrm{d}t,$$

这里 (以及下面) $\partial\vartheta$ 表示 ϑ 对应于前四个变量的次微分. 这样一来由 (7.13) 得

$$\left\{\begin{aligned}
&-x_{00}^* - x_{0N}^* - x_{0\,k+1}^* = \lambda x_0^* + \lambda h_N u_0^* + \lambda h_N y_0^* \\
&\qquad\qquad +2\lambda\big(\bar{x}_0 - \bar{x}(a)\big) - d_0^* - A^* d_N^*, \\
&-x_{jj}^* - x_{j\,j+N}^* = \lambda h_N u_j^* + \lambda h_N y_j^* - d_j^* - A^* d_{j+N}^*, \\
&\qquad\qquad j = 1, \cdots, k-N+1, \\
&-x_{jj}^* = \lambda h_N u_j^* - d_j^*, \quad j = k-N+2, \cdots, k, \\
&-x_{k+1\,k+1}^* = \lambda x_{k+1}^* - d_{k+1}^*, \\
&-z_{jj}^* = \lambda h_N z_j^* + d_j^* + e_{j-1}^* - e_j^*, \quad j = 0, \cdots, k, \\
&-v_{jj}^* = \lambda h_N v_j^* + \lambda \xi_j - h_N e_j^*, \quad j = 0, \cdots, k,
\end{aligned}\right. \tag{7.14}$$

其中

$$(x_0^*, x_{k+1}^*) \in \partial\varphi(\bar{x}_0^N, \bar{x}_{k+1}^N), \quad (u_j^*, y_{j-N}^*, z_j^*, v_j^*) \in \partial\vartheta(\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j^N, t_j),$$
$$\xi_j := 2\int_{t_j}^{t_{j+1}} \big(\bar{v}_j - \dot{\bar{z}}(t)\big)\,\mathrm{d}t,$$

这里在有穷维空间中不区别原空间和对偶空间中的元.

基于上面的关系, 就得到离散时间问题 (DA_N) 的如下必要最优条件, 其中

$$\theta_j(\cdot,\cdot,\cdot,\cdot) := \vartheta(\cdot,\cdot,\cdot,\cdot,t_j), \quad F_j(\cdot,\cdot,\cdot) := F(\cdot,\cdot,\cdot,t_j).$$

而对 F 的假设比 (D1) 和 (D2) 要弱一些, 但 (D4) 中 φ 和 ϑ 的连续性换成了 Lipschitz 连续性.

定理 7.3(差分–代数包含的必要最优条件) 设 $\bar{w}$ 为问题 (DA_N) 的一个最优解. 设 Ω 和 $\operatorname{gph} F_j$ 为局部闭, 函数 φ 和 ϑ_j 分别在 $(\bar{x}_0, \bar{x}_{k+1})$ 和 $(\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j)$ 附近 Lipschitz 连续 $(j = 0, \cdots, k)$. 则存在 $\lambda \geqslant 0$, $p_j \in \mathbb{R}^n$ $(j = 0, \cdots, k+N+1)$, $q_j \in \mathbb{R}^n$ $(j = -N, \cdots, k+1)$ 和 $r_j \in \mathbb{R}^n$ $(j = 0, \cdots, k+1)$, 不全为零, 满足条件

$$p_j = 0, \quad j = k+2, \cdots, k+N+1, \tag{7.15}$$

$$q_j = 0, \quad j = k-N+1, \cdots, k+1, \tag{7.16}$$

$$(p_0 + q_0, -p_{k+1}) \in \lambda\partial\varphi(\bar{x}_0, \bar{x}_{k+1}) + N\big((\bar{x}_0, \bar{x}_{k+1}); \Omega_N\big), \tag{7.17}$$

并且满足下面的与 Euler-Lagrange 包含类似的差分–代数包含:

$$\Big(\frac{\widetilde{p}_{j+1} - \widetilde{p}_j}{h_N}, \frac{\widetilde{q}_{j-N+1} - \widetilde{q}_{j-N}}{h_N}, \frac{r_{j+1} - r_j}{h_N}, -\frac{\lambda\xi_j}{h_N} + p_{j+1} + q_{j+1} + r_{j+1}\Big)$$
$$\in \lambda\partial\vartheta_j(\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j) + N\big((\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j, \bar{v}_j); \operatorname{gph} F_j\big),$$

其中 $j = 1, \cdots, k$, $\widetilde{p}_j := p_j + A^* p_{j+N}$, $\widetilde{q}_j := q_j + A^* q_{j+N}$.

证明 其实大部分的证明上面已经做了, 那里把 (MP) 的必要最优条件转换成了 (DA_N) 的在非光滑数学规划形式的对应条件. 下面需要做的只是改变 (7.14) 中的符号. 首先令

$$\widetilde{d}_j^* := \begin{cases} d_j^*, & j = 1, \cdots, k+1, \\ 0, & j = k+2, \cdots, k+N, \end{cases}$$

$$\widetilde{y}_j^* := \begin{cases} \lambda y_j^* + \dfrac{x_{j\,j+N}^*}{h_N}, & j = 1, \cdots, k-N+1, \\ 0, & j = k-N+2, \cdots, k, \end{cases}$$

以及 $\widetilde{r}_j := e_{j-1}^*$ $(j = 1, \cdots, k+1)$. 从 (7.14) 得关系

$$\left\{\begin{aligned}&\widetilde{d}_j^* + A^*\widetilde{d}_{j+N}^* - \widetilde{y}_j^* = \lambda u_j^* + \frac{x_{j\,j}^*}{h_N},\\&\widetilde{y}_{j-N}^* = \lambda y_{j-N}^* + \frac{x_{j-N\,j}^*}{h_N},\\&\frac{\widetilde{r}_{j+1} - \widetilde{r}_j}{h_N} - \widetilde{d}_j^* = \lambda z_j^* + \frac{z_{j\,j}^*}{h_N},\\&-\lambda\frac{\xi_j}{h_N} + \widetilde{r}_{j+1} = \lambda v_j^* + \frac{v_{j\,j}^*}{h_N}\end{aligned}\right. \tag{7.18}$$

对 $j = 1, \cdots, k$ 成立. 递归定义 $\widehat{p}_j$ 和 $\widehat{q}_j$ 如下:

$$\widehat{p}_j := \widehat{p}_{j+1} - h_N\widetilde{d}_j^*, \quad 而\ \widehat{p}_j = 0 \quad (j = k+2, \cdots, k+N+1),$$

$$\widehat{q}_j := \widehat{q}_{j+1} - h_N\widetilde{y}_j, \quad 而\ \widehat{q}_j = 0 \quad (j = k-N+1, \cdots, k+N+1).$$

置 $q_j := \widehat{q}_j + A^*\widehat{q}_{j+N}$, 则 (7.18) 可重写为

$$\left\{\begin{aligned}&\frac{(\widehat{p}_{j+1} - q_{j+1}) - (\widehat{p}_j - q_j)}{h_N} + A^*\frac{(\widehat{p}_{j+N+1} - q_{j+N+1}) - (\widehat{p}_{j+N} - q_{j+N})}{h_N}\\&= \lambda u_j^* + \frac{x_{j\,j}^*}{h_N}, \quad j = 1, \cdots, k,\\&\frac{(q_{j-N+1} + A^*q_{j+1}) - (q_{j-N} + A^*q_j)}{h_N} = \lambda y_{j-N}^* + \frac{x_{j-N\,j}^*}{h_N}, \quad j = 1, \cdots, k,\\&\frac{\widetilde{r}_{j+1} - \widetilde{r}_j}{h_N} - \frac{\widehat{p}_{j+1} - \widehat{p}_j}{h_N} = \lambda z_j^* + \frac{z_{j\,j}^*}{h_N}, \quad j = 1, \cdots, k,\\&-\lambda\frac{\xi_j}{h_N} + \widetilde{r}_{j+1} = \lambda v_j^* + \frac{v_{j\,j}^*}{h_N}, \quad j = 1, \cdots, k.\end{aligned}\right.$$

最后令

$$\begin{aligned}&p_0 := \lambda x_0^* + x_{0\,k+1}^* - q_0,\\&p_j := \widehat{p}_j - q_j, \quad j = 1, \cdots, k+N+1,\\&r_j := \widetilde{r}_j - \widehat{p}_j, \quad j = 1, \cdots, k+1.\end{aligned}$$

则得本定理中的必要最优条件. △

在附加的假设下, 下述推论证明的定理 7.3 的必要最优条件具有一些改进的非平凡性. 这在下一小节被用于连续时间问题 (DA) 最优条件的证明, 该证明基于对离散逼近取极限.

推论 7.4(差分–代数包含具有改进的非平凡性的必要条件) 在定理 7.3 的假设的基础上, 设映射 F_j $(j = 0, \cdots, k)$ 在 $(\bar{x}_j, \bar{x}_{j-N}, \bar{z}_j)$ 附近是局部有界和 Lipschitz 连续的. 则定理的必要条件成立且 $(\lambda, p_{k+1}, r_{k+1}) \neq 0$, 即可以要求下式成立:

$$\lambda^2 + \|p_{k+1}\|^2 + \|r_{k+1}\|^2 = 1. \tag{7.19}$$

证明 如果 $\lambda = 0$, 则定理的 Euler-Lagrange 包含除了蕴涵着条件 (7.15) 和 (7.16) 以外, 还有

$$\left(\frac{p_{k+1}-p_k}{h_N}, \frac{-q_{k-N}}{h_N}, \frac{r_{k+1}-r_k}{h_N}\right) \in D^*F_k(\bar{x}_k, \bar{x}_{k-N}, \bar{z}_k, \bar{v}_k)(-p_{k+1}-r_{k+1}).$$

若设 $p_{k+1}=0$ 和 $r_{k+1}=0$, 则

$$\left(\frac{-p_k}{h_N}, \frac{-q_{k-N}}{h_N}, \frac{-r_k}{h_N}\right) \in D^*F_k(\bar{x}_k, \bar{x}_{k-N}, \bar{z}_k, \bar{v}_k)(0).$$

根据推论 4.11 中有穷维空间上集值映射 Lipschitz 性质的上导数刻画, 可导出 $p_k = 0$, $q_{k-N}=0$ 和 $r_k=0$. 重复这个过程, 就得出和定理 7.3 的非平凡性表述相抵触的矛盾. △

7.1.4 微分–代数系统的 Euler-Lagrange 和 Hamilton 条件

本节的最后一小节推导由微分–代数包含控制的最优控制问题 (DA) 的广义 Euler-Lagrange 和 Hamilton 形式的必要最优条件. 这里从 Euler-Lagrange 条件开始, 它们在“松弛稳定性”的假设下给出了本节的主要结果. 下面定理中的符号 N_+ 和 ∂_+ 分别指的是 6.1.5 小节描述的移动结构的推广法锥和次微分. 与该小节中研究的常微分发展包含类似, 可以考虑具有“可加被积项”的问题 (DA), 从而可在 Euler-Lagrange 包含中用基本次微分 $\partial\vartheta$ 代替 $\partial_+\vartheta$.

定理 7.5(微分–代数包含的 Euler-Lagrange 条件) 设 $\{\bar{x}(\cdot), \bar{z}(\cdot)\}$ 为问题 (DA) 在假设 (D1)~(D4) 下的最优解, 其中假设中的函数 φ 和 $\vartheta(\cdot,\cdot,\cdot,\cdot,t)$ 的连续性对应地换成了 Lipschitz 连续性. 假设 (DA) 是松弛稳定的. 则存在 $\lambda \geqslant 0$, 逐段连续的轨线 $p\colon [a, b+\theta] \to \mathbb{R}^n$ 和 $q\colon [a-\theta, b] \to \mathbb{R}^n$ (仅可能在时滞 θ 倍数处可能有间断点), 以及绝对连续的轨线 $r\colon [a,b] \to \mathbb{R}^n$, 使得 $p(t) + A^*p(t+\theta)$ 和 $q(t-\theta)+A^*q(t)$ 在 $[a,b]$ 上绝对连续且满足关系

$$\lambda + \|p(b)\| + \|r(b)\| = 1, \tag{7.20}$$

$$p(t) = 0 \quad (\forall t \in (b, b+\theta]), \qquad q(t) = 0 \quad (\forall t \in (b-\theta, b]), \tag{7.21}$$

$$\big(p(a)+q(a), -p(b)\big) \in \lambda\partial\varphi\big(\bar{x}(a), \bar{x}(b)\big) + N\big((\bar{x}(a), \bar{x}(b)); \Omega\big) \tag{7.22}$$

和广义 Euler-Lagrange 包含

$$\begin{aligned}
&\left(\frac{\mathrm{d}}{\mathrm{d}t}\big[p(t)+A^*p(t+\theta)\big], \frac{\mathrm{d}}{\mathrm{d}t}\big[q(t-\theta)+A^*q(t)\big], \dot{r}(t)\right) \\
&\in \mathrm{co}\Big\{(u,v,w) \mid \big(u,v,w,p(t)+q(t)+r(t)\big) \in \lambda\widetilde{\partial}\vartheta\big(\bar{x}(t), \bar{x}(t-\theta), \bar{z}(t), \dot{\bar{z}}(t), t\big) \\
&\quad +N_+\big((\bar{x}(t), \bar{x}(t-\theta), \bar{z}(t), \dot{\bar{z}}(t)); \mathrm{gph}\, F(t)\big)\Big\}, \quad \text{a.e. } t \in [a,b].
\end{aligned}$$

证明 这里用离散逼近方法和前面得到的保证逼近问题 (DA_N) 的离散最优解和必要最优条件的强收敛性的结果. 为简化符号, 对 (DA_N) 的最优解和推论 7.4 中必要最优条件里的对应的元, 序号 N 都放在了上标里, 分别记为 $(\bar{x}^N, \bar{z}^N)$ 和 $(\lambda^N, p^N, q^N, r^N)$. 以 $\bar{x}^N(t)$, $p^N(t)$, $q^N(t-\theta)$ 和 $r^N(t)$ 表示这些离散轨线在 $[a,b]$ 上的逐段线性的拓展, 其对应的线性组合记为 $\bar{z}^N(t)$, $\widetilde{p}^N(t)$, 和 $\widetilde{q}^N(t-\theta)$. 由定理 7.1 得

$$\begin{aligned}\int_a^b \left\|\xi^N(t)\right\| \mathrm{d}t &= \sum_{j=0}^{k} \left\|\xi_j^N\right\| \leqslant 2\sum_{j=0}^{k} \int_{t_j}^{t_{j+1}} \left\|\dot{\bar{z}}(t) - \bar{v}_j^N\right\| \mathrm{d}t \\ &= 2\int_a^b \left\|\dot{\bar{z}}(t) - \dot{\bar{z}}^N(t)\right\| \mathrm{d}t := \nu_N \to 0 \quad (N \to \infty),\end{aligned}$$

其中 $\xi^N(t) := \xi_j^N / h_N$ $(t \in [t_j, t_{j+1}),\ j = 0, \cdots, k)$, $\xi_j^N = \xi_j$ 来自定理 7.3. 不失一般性, 设 $\lambda^N \to \lambda \geqslant 0$,

$$\bar{v}^N(t) := \dot{\bar{z}}^N(t) \to \dot{\bar{z}}(t), \quad \xi^N(t) \to 0, \quad \text{a.e. } t \in [a,b] \quad (N \to \infty).$$

下面对大的 N 估计 $\left(p^N(t), q^N(t-\theta), r^N(t)\right)$. 由 (7.15) 和 (7.16), 从定理 7.3 中的 Euler-Lagrange 包含可得

$$\begin{aligned}&\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N u_j^*, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N y_{j-N}^*, \frac{r_{j+1}^N - r_j^N}{h_N} - \lambda^N z_j^*, \\ &-\frac{\lambda^N \xi_j^N}{h_N} + p_{j+1}^N + r_{j+1}^N - \lambda^N v_j^*\Big) \in N\big((\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N, \bar{v}_j^N); \operatorname{gph} F_j\big)\end{aligned}$$

对某 $(u_j^*, y_{j-N}^*, z_j^*, v_j^*) \in \partial\vartheta_j(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N, \bar{v}_j^N)$ 成立 $(j = k - N + 2, \cdots, k+1)$. 根据上导数的定义, 这意味着

$$\begin{aligned}&\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N u_j^*, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N y_{j-N}^*, \frac{r_{j+1}^N - r_j^N}{h_N} - \lambda^N z_j^*\Big) \\ &\in D^* F_j\big(\bar{x}_j^N, \bar{x}_{j-N}^N, \bar{z}_j^N, \bar{v}_j^N\big)\Big(\lambda^N v_j^* + \frac{\lambda^N \xi_j^N}{h_N} - p_{j+1}^N - r_{j+1}^N\Big)\end{aligned}$$

对这个 j 成立. 由定理 4.11 中 Lipschitz 性质的上导数判据 (应用于具有模 ℓ_F 的 F_j) 得到估计

$$\begin{aligned}&\Big\|\Big(\frac{p_{j+1}^N - p_j^N}{h_N} - \lambda^N u_j^*, \frac{q_{j-N+1}^N - q_{j-N}^N}{h_N} - \lambda^N y_{j-N}^*, \frac{r_{j+1}^N - r_j^N}{h_N} - \lambda^N z_j^*\Big)\Big\| \\ &\leqslant \ell_F \Big\|\lambda^N v_j^* + \frac{\lambda^N \xi_j^N}{h_N} - p_{j+1}^N - r_{j+1}^N\Big\| \quad (j = k - N + 2, \cdots, k+1).\end{aligned}$$

因为 ϑ 为 Lipschitz 连续且具有模 ℓ_ϑ, 所以 $\|(u_j^*, y_{j-N}^*, z_j^*, v_j^*)\| \leqslant \ell_\vartheta$, 则由上述得

$$
\begin{aligned}
&\|(p_j^N, q_{j-N}^N, r_j^N)\| \\
\leqslant\,& \ell_F\|\xi_j^N\| + (\ell_F+1)h_N\ell_\vartheta + (\ell_F h_N+1)\|(p_{j+1}^N, q_{j-N+1}^N, r_{j+1}^N)\| \\
\leqslant\,& \ell_F\|\xi_j^N\| + (\ell_F h_N+1)\ell_F\|\xi_{j+1}^N\| + (\ell_F+1)h_N\ell_\vartheta + (\ell_F h_N+1)(\ell_F+1)h_N\ell_\vartheta \\
&+(\ell_F h_N+1)^2\|(p_{j+2}^N, q_{j-N+2}^N, r_{j+2}^N\| \leqslant \cdots \\
\leqslant\,& \exp\big(\ell_F(b-a)\big)\Big(1+\frac{\ell_\vartheta}{\ell_F}(\ell_F+1)+\ell_F\nu_N\Big) \quad (j=k-N+2,\cdots,k+1).
\end{aligned}
$$

这蕴涵着 $\{(p_j^N, q_{j-N}^N, r_j^N) \mid j = k-N+2,\cdots,k+1\}$ 有界, 从而也有 $(p^N(t), q^N(t-\theta), r^N(t))$ 在 $[b-\theta, b]$ 上一致有界.

接下来考虑下标 $j=k-2N+2,\cdots,k-N+1$. 从 Euler-Lagrange 包含得

$$
\begin{aligned}
&\Big\|\Big(\frac{p_{j+1}^N-p_j^N}{h_N}-\lambda^N u_j^*, \frac{q_{j-N+1}^N-q_{j-N}^N}{h_N}-\lambda^N y_{j-N}^*, \frac{r_{j+1}^N-r_j^N}{h_N}-\lambda^N z_j^*\Big)\Big\| \\
\leqslant\,& \ell_F\Big\|\lambda^N v_j^* + \frac{\lambda^N\xi_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N - r_{j+1}^N\Big\| \\
&+\Big\|\Big(\frac{A^*p_{j+N+1}^N - A^*p_{j+N}^N}{h_N}, \frac{A^*q_{j+1}^N - A^*q_j^N}{h_N}, 0\Big)\Big\|.
\end{aligned}
$$

根据前面提到的上导数判据和 p_j^N 与 q_j^N 的一致有界性 (界限为某 $\alpha>0$), 可得

$$
\begin{aligned}
&\Big\|\Big(\frac{p_{j+1}^N-p_j^N}{h_N}-\lambda^N u_j^*, \frac{q_{j-N+1}^N-q_{j-N}^N}{h_N}-\lambda^N y_{j-N}^*, \frac{r_{j+1}^N-r_j^N}{h_N}-\lambda^N z_j^*\Big)\Big\| \\
\leqslant\,& \ell_F\Big\|\lambda^N v_j^* + \frac{\lambda^N\xi_j^N}{h_N} - p_{j+1}^N - q_{j+1}^N - r_{j+1}^N\Big\| + \frac{\alpha}{h_N}
\end{aligned}
$$

对 $j=k-2N+2,\cdots,k-N+1$ 成立. 于是有估计

$$
\begin{aligned}
&\|(p_j^N, q_{j-N}^N, r_j^N)\| \\
\leqslant\,& \ell_F\|\xi_j^N\| + (\ell_F+1)h_N\ell_\vartheta + (\ell_F h_N+1)\|(p_{j+1}^N, q_{j-N+1}^N, r_{j+1}^N)\| + (\ell_F h_N+1)\alpha \\
\leqslant\,& \ell_F\|\xi_j^N\| + (\ell_F h_N+1)\ell_F\|\xi_{j+1}^N\| + (\ell_F+1)h_N\ell_\vartheta + (\ell_F h_N+1)(\ell_F+1)h_N\ell_\vartheta \\
&+(\ell_F h_N+1)(\ell_F+1)\alpha + (\ell_F h_N+1)^2\|(p_{j+2}^N, q_{j-N+2}^N, r_{j+2}^N\| \\
\leqslant\,& \cdots \leqslant \exp\big(\ell_F(b-a)\big)\Big(1+\frac{(\ell_\vartheta+\alpha)(\ell_F+1)}{\ell_F}+\ell_F\nu_N\Big) \\
&(j=k-2N+2,\cdots,k-N+1).
\end{aligned}
$$

这蕴涵着 p_j^N, q_{j-N}^N 和 r_j^N 对 $j=k-2N+2,\cdots,k-N+1$ 是有界的, 从而 $\{p^N(t), q^N(t-\theta), r^N(t)\}$ 在 $[b-2\theta, b-\theta]$ 上是一致有界的. 重复这个过程, 可得 $\{p^N(t), q^N(t-\theta), r^N(t)\}$ 和 $\{\widetilde{p}^N(t), \widetilde{q}^N(t-\theta)\}$ 在 $[a,b]$ 都是一致有界的.

应用离散 Euler-Lagrange 包含和局部 Lipschitz 性质的上导数判据, 下面在 $[a,b]$ 上估计 $(\dot{\widetilde{p}}^N(t),\dot{\widetilde{q}}^N(t-\theta),\dot{r}^N(t))$. 对 $t_j\leqslant t\leqslant t_{j+1}$ $(j=0,\cdots,k)$, 有

$$\begin{aligned}
&\left\|\left(\dot{\widetilde{p}}^N(t),\dot{\widetilde{q}}^N(t-\theta),\dot{r}^N(t)\right)\right\|\\
=&\left\|\left(\frac{\widetilde{p}_{j+1}-\widetilde{p}_j}{h_N},\frac{\widetilde{q}_{j-N+1}-\widetilde{q}_{j-N}}{h_N},\frac{r^N_{j+1}-r^N_j}{h_N}\right)\right\|\\
\leqslant&\ell_F\left\|\lambda^N v_j^*+\frac{\lambda^N\xi_j^N}{h_N}-p^N_{j+1}-q^N_{j+1}-r^N_{j+1}\right\|+\ell_\vartheta\\
\leqslant&\ell_F\|\xi^N\|+\ell_F\|p^N_{j+1}\|+\ell_F\|q^N_{j+1}\|+\ell_F\|r^N_{j+1}\|+(\ell_F+1)\ell_\vartheta.
\end{aligned}$$

于是序列 $\{\dot{\widetilde{p}}^N(t),\dot{\widetilde{q}}^N(t-\theta),\dot{r}^N(t)\}$ 在 $L^N[a,b]$ 中是弱紧的. 则可找到三个在 $[a,b]$ 上绝对连续的映射 $\widetilde{p}(\cdot)$, $\widetilde{q}(\cdot-\theta)$ 和 $r(\cdot)$ 使得在 $L^1[a,b]$ 的弱拓扑下有

$$\dot{\widetilde{p}}^N(t)\to\dot{\widetilde{p}}(t),\quad \dot{\widetilde{q}}^N(t-\theta)\to\dot{\widetilde{q}}(t-\theta),\quad \dot{r}^N(t)\to\dot{r}(t),$$

并且 $\widetilde{p}^N(t)\to\widetilde{p}(t)$, $\widetilde{q}^N(t-\theta)\to\widetilde{q}(t-\theta)$, $r^N(t)\to r(t)$ 在 $[a,b]$ 上一致成立 (其中 $N\to\infty$; 不失一般性, 这里设这些序列都是整个序列收敛). 因为 $p^N(t)$ 和 $q^N(t-\theta)$ 在 $[a,b+\theta]$ 上一致有界, 则它们必然在 $L^1[a,b+\theta]$ 的弱拓扑下分别收敛到某轨线 $p(t)$ 和 $q(t-\theta)$. 考虑上面 $\widetilde{p}^N(t)$ 和 $\widetilde{q}^N(t-\theta)$ 的收敛性, 由 (7.16) 得 $p(\cdot)$ 和 $q(\cdot)$ 满足 (7.21),

$$\widetilde{p}(t)=p(t)+A^*p(t+\theta),\quad \widetilde{q}(t-\theta)=q(t-\theta)+A^*q(t)\quad (t\in[a,b]),$$

且 $p(t)$ 和 $q(t)$ 分别在 $[a,b+\theta]$ 和 $[a-\theta,b]$ 上分段连续, 仅可能在点 $b-i\theta$ $(i=0,1,\cdots)$ 有间断 (即右方不连续). 由有穷维空间中基本法锥和次微分的鲁棒性质, 分别对 (7.19) 和 (7.17) 取极限可得条件 (7.20) 和 (7.22).

剩下需要证明定理中的广义 Euler-Lagrange 包含. 为此重写定理 7.3 中的离散 Euler-Lagrange 包含为如下形式

$$\begin{aligned}
&\left(\dot{\widetilde{p}}^N(t),\dot{\widetilde{q}}^N(t-\theta),\dot{r}^N(t)\right)\\
\in&\Big\{(u,v,w)\mid\Big(u,v,w,p^N(t_{j+1})+q^N(t_{j+1})+r^N(t_{j+1})-\frac{\lambda^N\xi_j^N}{h_N}\Big)\\
&\in\lambda^N\partial\vartheta\left(\bar{x}^N(t_j),\bar{x}^N(t_j-\theta),\bar{z}^N(t_j),\bar{v}_j^N\right)\\
&+N\left((\bar{x}^N(t_j),\bar{x}^N(t_j-\theta),\bar{z}^N(t_j),\bar{v}_j^N);\operatorname{gph}F(t_j)\right)\Big\},
\end{aligned}\tag{7.23}$$

其中 $t\in[t_j,t_{j+1}]$, $j=0,\cdots,k$. 从经典的 Mazur 定理得到 $(\dot{\widetilde{p}}^N(t),\dot{\widetilde{q}}^N(t-\theta),\dot{r}^N(t))$ 的凸组合组成的序列收敛到 $(\dot{\widetilde{p}}(t),\dot{\widetilde{q}}(t-\theta),\dot{r}(t))$ 对 $t\in[a,b]$ 几乎处处成立. 在 (7.23) 中取极限, 并考虑到上面建立的 $\xi^N(t)$ 和 $\bar{v}^N(t)$ 的逐点收敛性、扩展法锥和

次微分的构造及它们对所有变量和参数的鲁棒性质, 就得到欲证的问题 (DA) 的 Euler-Lagrange 包含. 证毕. △

可以看到, 对 Mayer 问题 (DA_M) (即在 $\vartheta = 0$ 时的 (DA)), 定理 7.5 中推广的 Euler-Lagrange 包含可以等价地写成

$$\Big(\frac{\mathrm{d}}{\mathrm{d}t}\big[p(t) + A^*p(t+\theta)\big], \frac{\mathrm{d}}{\mathrm{d}t}\big[q(t-\theta) + A^*q(t)\big], \dot r(t)\Big)$$
$$\in \mathrm{co}\, D_+^*F\big(\bar x(t), \bar x(t-\theta), \bar z(t), \dot{\bar z}(t)\big)\big(-p(t) - q(t) - r(t)\big), \quad \text{a.e. } t \in [a,b],$$

这里 F 的扩展上导数是对变量 (x, y, z) 取的, $t \in [a, b]$ 是移动参数, 其中移动 (对参数 $t \in T$) 映射 $S\colon X \times T \rightrightarrows Y$ 在 $(\bar x, \bar y, \bar t)$ (其中 $\bar y \in S(\bar x, \bar t)$) 的扩展上导数定义为

$$D_+^*S(\bar x, \bar y, \bar t)(y^*) := \big\{x^* \in X^* \mid (x^*, -y^*) \in N_+\big((\bar x, \bar y); \mathrm{gph}\, S(\cdot, \bar t)\big)\big\}, \quad y^* \in Y^*.$$

可以证明, 上面得到的广义 Euler-Lagrange 包含在原问题的松弛稳定性假设下蕴涵另外两个关键的最优条件, 它们由基于速度映射 F 的 Hamilton 函数表述. 第一个条件称为"扩展 Hamilton 包含", 在下面以 Hamilton 函数基本次微分的部分凸化给出; 第二个条件是所考虑的微分–代数包含的经典的 Weierstrass-Pontryagin 最大值条件的类似结果. 我们知道, 最大值原理 (极大条件的核心) 的类似结果一般来说对微分–代数系统是不成立的, 即使在由中立型光滑泛函微分方程控制的最优控制问题上也是这样, 而这个问题是 (DA) 的一个特殊情况.

在有穷维的常微分包含的情形 (比照注 6.32), 下述关于广义 Euler-Lagrange 和 Hamilton 包含的关系基于 Rockafellar 的对偶化定理, 该定理涉及 (无论有无动态关系) 相关于集值映射的抽象 Lagrange 和 Hamilton 函数的次导数. 简单起见, 这里只考虑自治微分–代数系统的 Mayer 问题 (DA_M). 此时映射 F 的 Hamilton 函数定义为

$$\mathcal{H}(x, y, z, p) := \sup\big\{\langle p, v\rangle \mid v \in F(x, y, z)\big\}.$$

推论 7.6(微分–代数包含的扩展 Hamilton 包含和最大值条件)　设 $\{\bar x(\cdot), \bar z(\cdot)\}$ 为自治时滞微分–代数系统在定理 7.5 假设下的 Mayer 问题 (DA_M) 的一个最优解, 则存在一个常数 $\lambda \geqslant 0$, 逐段连续的轨线 $p\colon [a, b+\theta] \to \mathbb{R}^n$ 和 $q\colon [a-\theta, b] \to \mathbb{R}^n$ (仅可能在时滞 θ 的倍数处有间断), 绝对连续的轨线 $r\colon [a,b] \to \mathbb{R}^n$, 使得 $p(t) + A^*p(t+\theta)$ 和 $q(t-\theta) + A^*q(t)$ 在 $[a, b]$ 上绝对连续, 且在 (7.20)~(7.22) 之外, 还成立如下的扩展 Hamilton 包含

$$\begin{aligned}&\Big(\frac{\mathrm{d}}{\mathrm{d}t}\big[p(t) + A^*p(t+\theta)\big], \frac{\mathrm{d}}{\mathrm{d}t}\big[q(t-\theta) + A^*q(t)\big], \dot r(t)\Big)\\ &\in \mathrm{co}\Big\{(u, v, w) \mid \big(-u, -v, -w, \dot{\bar z}(t)\big)\\ &\quad \in \partial\mathcal{H}\big(\bar x(t), \bar x(t-\theta), \bar z(t), p(t) + q(t) + r(t)\big)\Big\}\end{aligned} \tag{7.24}$$

和最大值条件

$$\langle p(t)+q(t)+r(t),\dot{\bar z}(t)\rangle=\mathcal{H}\big(\bar x(t),\bar x(t-\theta),\bar z(t),p(t)+q(t)+r(t)\big) \tag{7.25}$$

对 $t\in[a,b]$ 几乎处处成立. 如果进一步 F 在 $\big(\bar x(t),\bar x(t-\theta),\bar z(t)\big)$ 附近是凸值的, 则 (7.24) 等价于 Euler-Lagrange 包含

$$\begin{aligned}&\Big(\frac{\mathrm{d}}{\mathrm{d}t}\big[p(t)+A^*p(t+\theta)\big],\frac{\mathrm{d}}{\mathrm{d}t}\big[q(t-\theta)+A^*q(t)\big),\dot r(t)\Big)\\&\in\operatorname{co}D^*F\big(\bar x(t),\bar x(t-\theta),\bar z(t),\dot{\bar z}(t)\big)\big(-p(t)-q(t)-r(t)\big)\end{aligned} \tag{7.26}$$

对 $t\in[a,b]$ 几乎处处成立, 在所考虑的情形这自动蕴涵最大值条件 (7.25).

证明 因为 (DA_M) 对松弛是稳定的, 二元组 $\{\bar x(\cdot),\bar z(\cdot)\}$ 是松弛问题 $(\overline{\mathrm{DA}}_M)$ 的一个最优解, 而 $(\overline{\mathrm{DA}}_M)$ 与 (DA_M) 唯一的区别是原来的时滞微分–代数包含换成了其凸化 (7.10). 根据定理 7.5, 最优解 $\{\bar x(\cdot),\bar z(\cdot)\}$ 满足条件 (7.20)~(7.22) 以及 Euler-Lagrange 包含 (7.26) 的松弛版本, 其中的 F 替换成了其凸包 $\operatorname{co}F$. 由 Rockafellar 的对偶化定理得

$$\begin{aligned}&\operatorname{co}\Big\{(u,v,w)\mid(u,v,w,p)\in N\big((x,y,z,q);\operatorname{gph}(\operatorname{co}F)\big)\Big\}\\=&\operatorname{co}\Big\{(u,v,w)\mid(-u,-v,-w,q)\in\partial\overline{\mathcal{H}}(x,y,z,p)\Big\},\end{aligned}$$

其中 $\overline{\mathcal{H}}$ 表示松弛系统的 Hamilton 函数, 即将 $\mathcal{H}$ 中的 F 换为 $\operatorname{co}F$. 容易验证 $\overline{\mathcal{H}}=\mathcal{H}$. 这样松弛系统的扩展 Euler-Lagrange 包含就蕴涵扩展的 Hamilton 包含 (7.24), 它当然就蕴涵最大值条件 (7.25). 若 F 为凸值, 由上述对偶化等式知 (7.24) 和 (7.26) 等价. 由定理 1.34 可注意到, F 为凸值时 Euler-Lagrange 包含 (7.26) 蕴涵这最大值条件 (7.25). 这对松弛稳定性也对, 其中松弛问题中的伴随轨线 $\big(p(\cdot),q(\cdot),r(\cdot)\big)$ 满足 Euler-Lagrange 包含. △

注 7.7(时滞微分包含的最优控制) 当最优控制问题由

$$\dot x(t)\in F\big(x(t),x(t-\theta),t\big),\quad \text{a.e. } t\in[a,b]$$

(即仅在状态变量上有时滞) 类型的“时滞微分包含”控制, 则所得结果可以更具体和简化. 与本节研究的微分–代数和中立型系统相比, 这种系统更接近于常微分包含问题. 而和常微分/微分–代数/中立型相比, 该系统具有的一个值得一提的特殊性质是, 当初始时间区间上具有集值尾部约束

$$x(t)\in C(t),\quad \text{a.e. } t\in[a-\theta,a)$$

时可以得到很有价值的结果, 该约束为优化提供了额外的来源; 更多细节见 Mordukhovich 和 L. Wang[973]. 进一步, 6.2 节在没有松弛假设的情况下建立的逼近过

程和必要最优条件可以扩展到时滞微分系统, 与常微分发展包含相比, 并不需要很大的改动.

但是, 对微分–代数和中立型的包含 (即在 (DA) 中 $A \neq 0$ 的情形) 似乎得不到类似的最优条件. 这里的主要原因是, 6.3 节中的逼近过程及结果本质上是基于自由端点具有有限积分项的 Bolza 问题自动成立的松弛稳定性, 而对速度变量具有时滞或状态变量具有代数关系的问题, 这并不成立.

7.2 半线性约束双曲方程的 Neumann 边界控制

本节研究一类半线性双曲方程的最优控制问题, 其控制作用于 Neumann 边界条件上, 控制和状态函数具有逐点约束. 众所周知, 状态约束控制问题在动态优化中是最困难且具有挑战性的问题之一. 尽管这样的问题在常微分和时滞控制系统以及椭圆和抛物型偏微分的情形有广泛的研究, 对双曲方程却非如此. 而且, 边界控制问题从本质上复杂于控制参数出现在微分方程中的情形 (即涉及所谓分布控制的问题).

本节研究具有状态约束双曲系统的 Neumann 边界控制问题, 对应的控制在 Dirichlet 边界条件的问题 (与 Neumann 类型本质上是不同的) 下一节讨论. 主要目标是对一个受控于半线性波动方程的状态约束 Neumann 边界问题建立必要最优条件, 其具有逐点最大值原理的形式, 当中的假设是适度和自然的. 所用方法基于变分分析中的扰动方法, 它涉及状态约束的罚函数以及对无约束逼近问题的必要条件取极限; 比照 6.2 节发展包含的情形. 但本节对无约束双曲方程的逼近控制问题的分析与 6.2 节是不同的, 它是基于与 6.3 节常微分控制系统类似的针形变分的. 细节见下.

7.2.1 问题的表述和 Neumann 边界控制的必要最优条件

给定具有 $\mathcal{C}^2$ 边界 Γ 的有界开集 (定义域) $\Omega \subset \mathbb{R}^n$ 和正数 (时间) T, 这里主要讨论下述受控于半线性波动方程的最优控制问题:

$$\min\ J(y,u) = \int_\Omega f\big(x, y(T)\big)\,\mathrm{d}x + \int_Q g(x,t,y)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma h(s,t,u)\,\mathrm{d}s\,\mathrm{d}t,$$

其中容许解对 $\{y(\cdot), u(\cdot)\}$ 满足

$$\begin{cases} y_{tt} - \Delta y + \vartheta(\cdot, y) = 0, & \text{在 } Q := \Omega \times (0,T) \text{ 中}, \\ \partial_\nu y = u, & \text{在 } \Sigma := \Gamma \times (0,T) \text{ 中}, \\ y(0) = y_0, \quad y_t(0) = y_1, & \text{在 } \Omega \text{ 中}, \end{cases} \tag{7.27}$$

并具有逐点的控制和状态函数约束

$$u(\cdot) \in U_{ad} \subset L^2(\Sigma), \quad y(\cdot) \in \Theta \subset \mathcal{C}\big([0,T]; L^2(\Omega)\big),$$

其中 Δ 表示经典的 Laplace 算子, ∂_ν 表示边界上的通常法向导数. 记该问题为 (NP) 并简写为

$$\inf\Big\{ J(y,u) \mid \{y(\cdot),u(\cdot)\}\text{满足}(7.27), u(\cdot) \in U_{ad}, y(\cdot) \in \Theta \Big\}.$$

对非线性函数 ϑ 和被积项 f, g 与 h 的假设见下. 初始状态 $(y_0, y_1) \in H^1(\Omega) \times L^2(\Omega)$ 是固定的. 需要指出, 本节的主要构造和结果可以推广到受控于 (7.27) 中更一般的强椭圆算子 (不仅仅是 Laplace 算子 Δ) 的双曲方程, 其具有不依赖于时间且正则 (在偏微分方程的通常意义下) 的系数.

本节和后面的各节都使用偏微分控制文献中的标准记号. 方便读者, 重述如下. 集合 $\mathcal{M}\big([0,T]; L^2(\Omega)\big)$ 指 $[0,T]$ 上取值于 $L^2(\Omega)$ 中测度的空间, 它是 $\mathcal{C}\big([0,T]; L^2(\Omega)\big)$ 的拓扑对偶. 空间

$$\mathcal{C}_0\big(]0,T]; L^2(\Omega)\big) := \big\{ y \in \mathcal{C}\big([0,T]; L^2(\Omega)\big) \mid y(0) = 0 \big\}$$

的拓扑对偶记为 $\mathcal{M}_b\big(]0,T]; L^2(\Omega)\big)$. 类似地, 空间

$$\mathcal{C}_0\big(]0,T[; L^2(\Omega)\big) := \big\{ y \in \mathcal{C}\big([0,T]; L^2(\Omega)\big) \mid y(0) = 0,\ y(T) = 0 \big\}$$

的拓扑对偶记为 $\mathcal{M}_b\big(]0,T[; L^2(\Omega)\big)$. 注意空间 $\mathcal{C}_0\big(]0,T]; L^2(\Omega)\big)$ 和 $\mathcal{C}_0\big(]0,T[; L^2(\Omega)\big)$ 由闭区间 $[0,T]$ 上的连续映射组成, 其中的一边或双边的端点是给定的. 在下面的表述中, $]0,T]$ 和 $]0,T[$ 分别与 $(0,T]$ 和 $(0,T)$ 是一样的.

众所周知, 任何测度 $\mu \in \mathcal{M}\big(]0,T]; L^2(\Omega)\big)$ 等同于一个满足 $\widetilde{\mu}(\{0\}) = 0$ 和 $\widetilde{\mu}|_{]0,T]} = \mu$ 的测度 $\widetilde{\mu} \in \mathcal{M}\big([0,T]; L^2(\Omega)\big)$, 其中 $\widetilde{\mu}|_{]0,T]}$ 表示限制在 $]0,T]$ 上的 $\widetilde{\mu}$. 因此若 $y \in \mathcal{C}\big([0,T]; L^2(\Omega)\big)$ 且 $\mu \in \mathcal{M}_b\big(]0,T]; L^2(\Omega)\big)$, 这里仍然以

$$\langle y, \mu \rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T];L^2(\Omega))} \quad \text{表示} \quad \langle y, \widetilde{\mu} \rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}([0,T];L^2(\Omega))}.$$

因为要处理在 Q 上的分布意义下得到满足的方程, 把 $\mathcal{M}_b\big(]0,T]; L^2(\Omega)\big)$ 对等于 $\mathcal{M}_b\big(\Omega \times]0,T]\big)$ 的一个子空间是方便的, 这源于连续稠密嵌入 $\mathcal{C}_0\big(\Omega \times]0,T]\big) \hookrightarrow \mathcal{C}_0\big(]0,T]; L^2(\Omega)\big)$. 所以对 $\mu \in \mathcal{M}_b\big(]0,T]; L^2(\Omega)\big)$, 记号 $\mu|_Q$ (即 μ 在 Q 上的限制) 在 μ 被看成是 $\Omega \times]0,T] = \Omega \times (0,T]$ 上的有界测度时是有意义的, 因此 $\mu|_{\Omega \times \{T\}}$ 表示 $\mu(\{T\})$. 类似情形下面用了相同的记号. 对 $z \in L^2(Q)$, 用 z_t (对应地, z_{tt}) 表示 z 对 t 在 Q 中分布意义下的导数 (对应地, 二阶导数).

给定 Banach 空间 Z, Z 和 Z^* 的对偶对记为 $\langle \cdot, \cdot \rangle_{Z,Z^*}$. 没有混淆的情形则以 $\langle \cdot, \cdot \rangle$ 替代 $\langle \cdot, \cdot \rangle_{Z,Z^*}$. 在需要强调所考虑的双曲方程解的某种特殊正则性时, 写法略

有不同, 比如说 $(y, y_t) \in \mathcal{C}([0,T];X) \times \mathcal{C}([0,T];Y)$ 是 (7.27) 的解, 而不是仅指出 y 是该系统的解.

若 $p(\cdot)$ 属于 $BV([0,T];H^1(\Omega)^*)$ (即 $[0,T]$ 上取值于 $H^1(\Omega)^*$ 的有界变差函数空间), 可以对任意 $t \in (0,T)$ 定义 $p(t^-)$ 和 $p(t^+)$, 以及 $p(0^+)$ 和 $p(T^-)$, 其中 $p(0)$ 和 $p(T)$ 的值与 $p(0^+)$ 和 $p(T^-)$ 一般是可以不同的. 在 $[0,T]$ 上取值于 $H^1(\Omega)^*$ 有唯一 Radon 测度, 记为 d_tp, 使得 d_tp 在 $(0,T)$ 上的限制是 p 在 $(0,T)$ 上的向量值分布意义下的导数且满足 $d_tp(\{0\}) = p(0^+) - p(0)$ 和 $d_tp(\{T\}) = p(T) - p(T^-)$. 进一步, 把 p 等同于在 $(0,T)$ 上右方连续的表示, 则

$$p(0^+) = p(0) + d_tp(\{0\}), \quad p(t) = p(0) + d_tp([0,t]), \quad \forall t \in]0,T].$$

若 $\{p_k\}$ 是 $BV([0,T];H^1(\Omega)^*)$ 中的有界序列, 则存在子序列 $\{p_{k_m}\}$ 和函数 $p \in BV([0,T];H^1(\Omega)^*)$ 满足, 对几乎所有的 $t \in [0,T]$, 在 $H^1(\Omega)^*$ 的弱拓扑下有

$$p_{k_m}(t) \to p(t).$$

注意若没有具体给定上面的在 $(0,T)$ 上右方连续的表示, 则该收敛性可能对任意 $t \in [0,T]$ 都成立, 比如可见 Barbu 与 Precupanu [84]. 特别地, 沿 $H^1(\Omega)^*$ 的弱拓扑在 $m \to \infty$ 时有

$$p_{k_m}(T) \to p(T).$$

现在描述问题 (NP) 初始数据上的常驻假设.

(H1) 对任意 $y \in \mathbb{R}$, 函数 $\vartheta(\cdot,\cdot,y)$ 在 Q 上可测; 对几乎处处的 $(x,t) \in Q$, 函数 $\vartheta(x,t,\cdot)$ 为 $\mathcal{C}^1$ 的. 进一步, 有

$$\vartheta(\cdot,0) \in L^1(0,T;L^2(\Omega)), \quad |\vartheta'_y(x,t,y)| \leqslant M$$

在 $Q \times \mathbb{R}$ 中成立, 其中 $M > 0$, ϑ'_y 表示偏导数.

(H2) 对任意 $y \in \mathbb{R}$, 函数 $f(\cdot,y)$ 在 Ω 上可测且 $f(\cdot,0) \in L^1(\Omega)$. 对几乎处处的 $x \in \Omega$, $f(x,\cdot)$ 是 $\mathcal{C}^1$ 的. 进一步, 存在常数 $C > 0$ 满足

$$|f'_y(x,y)| \leqslant C(1+|y|), \quad 当 (x,y) \in \Omega \times \mathbb{R}.$$

(H3) 对任意 $y \in \mathbb{R}$, 函数 $g(\cdot,\cdot,y)$ 在 Q 上可测且 $g(\cdot,0) \in L^1(Q)$. 对几乎处处的 $(x,t) \in Q$, $g(x,t,\cdot)$ 为 $\mathcal{C}^1$ 的. 进一步, 存在常数 $C > 0$ 满足

$$|g'_y(x,t,y)| \leqslant C(1+|y|), \quad 当 (x,t,y) \in Q \times \mathbb{R}.$$

(H4) 对任意 $u \in \mathbb{R}$, 函数 $h(\cdot,\cdot,u)$ 在 Σ 上可测且 $h(\cdot,0) \in L^1(\Sigma)$. 对几乎处处的 $(s,t) \in \Sigma$, $h(s,t,\cdot)$ 为 $\mathcal{C}^1$ 的. 进一步, 存在 $C > 0$ 满足

$$|h'_u(s,t,u)| \leqslant C(1+|u|), \quad 当 (s,t,u) \in \Sigma \times \mathbb{R}.$$

(**H5**) 状态约束集 $\Theta \subset \mathcal{C}\big([0,T];L^2(\Omega)\big)$ 是闭凸的且 $\operatorname{int}\Theta \neq \varnothing$. 进一步假设初始状态函数 $\widehat{y}_0(x,t) := y_0(x)$ 属于 Θ 的内部.

(**H6**) 控制集 U_{ad} 由如下形式给出:

$$U_{ad} := \big\{u \in L^2(\Sigma) \mid u(s,t) \in K(s,t), \quad \text{a.e. } (s,t) \in \Sigma\big\},$$

其中 $K(\cdot)$ 为可测多值函数, 其值为 $\mathbb{R}$ 中的非空闭集.

当然还假设 (NP) 的可行解对集合 $\{y(\cdot),u(\cdot)\}$ 非空, 即存在 $u(\cdot) \in U_{ad}$ 满足 $J(y,u) < \infty$, 其中 $y(\cdot) \in \Theta$ 是系统 (7.27) 对应于 u 的弱解; 精确定义见下一小节.

注意到, 上面的基本假设对费用函数中被积项 (对状态或控制变量) 和控制集 U_{ad} 没有任何凸性要求, 这与 7.3 节中讨论的 Dirichlet 边界控制的假设是不同的. 原因在于, Neumann 边值问题具有更多的正则性, 从而可以应用有力的变分方法不依赖于弱收敛来证明最优条件; 更多讨论见 7.3 节.

为表述本节的主要结果, 对控制问题 (NP) 定义 Hamilton-Pontryagin 函数 (的类似版本):

$$H(s,t,u,p,\lambda) := pu + \lambda h(s,t,u).$$

下述定理给出了 (NP) 最优解的必要条件, 它提供了所考虑的 Neumann 边界控制问题的一个逐点形式的 Pontryagin 最大值原理. 这里表述成最小值 (而不是最大值) 条件更为方便一些.

定理 7.8(Neumann 边界控制的逐点必要最优条件) 设 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$ 为问题 (NP) 满足假设 (H1)~(H6) 的最优解. 则存在 $\lambda \geqslant 0$, $\mu \in \mathcal{M}\big(]0,T];L^2(\Omega)\big)$, 和可测集 $\widetilde{\Sigma} \subset \Sigma$ 使得 $\mathcal{L}^n(\Sigma\backslash\widetilde{\Sigma}) = 0$,

$$(\lambda,\mu) \neq 0, \quad \langle \mu, z-\bar{y}\rangle \leqslant 0, \quad \forall z \in \Theta, \tag{7.28}$$

$$H\big(s,t,\bar{u}(s,t),p(s,t),\lambda\big) = \min_{u\in K(s,t)} H\big(s,t,u,p(s,t),\lambda\big) \tag{7.29}$$

对任意 $(s,t) \in \widetilde{\Sigma}$ 成立, 其中 $\mathcal{L}^n$ 表示 n 维 Lebesgue 测度, $p(\cdot)$ 是对应的如下伴随系统的解:

$$\begin{cases} p_{tt} - \Delta p + \vartheta'_y(\cdot,\bar{y})p = \lambda g'_y(x,t,\bar{y}) + \mu|_Q, & \text{在 } Q \text{ 中}, \\ \partial_\nu p = 0, & \text{在 } \Sigma \text{ 中}, \\ p(T) = y_0, \quad p_t(T) = -\lambda f'_y\big(x,\bar{y}(T)\big) - \mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中}. \end{cases} \tag{7.30}$$

定理 7.8 的证明在 7.2.4 小节给出. 状态和伴随系统解的定义在 7.2.2 小节中表述和讨论.

7.2.2 Neumann 问题中状态和伴随系统的分析

首先考虑线性波动方程经典的非齐次 Neumann 边值问题

$$\begin{cases} y_{tt} - \Delta y = \phi, & \text{在 } Q \text{ 中}, \\ \partial_\nu y = u, & \text{在 } \Sigma \text{ 中}, \\ y(0) = y_0, \quad y_t(0) = y_1, & \text{在 } \Omega \text{ 中}. \end{cases} \tag{7.31}$$

下面的基本正则性结果由 Lasiecka 与 Triggiani[744, 745] 建立; 其 (很难的) 证明、讨论及在偏微分方程上的应用建议读者参考原文. 这里的目标是在变分分析的框架下把这个结果集成到所考虑的 Neumann 边界控制问题中. 这个分析中的一个很有意义的部分在本节给出, 该研究涉及了具有 Neumann 边界控制的双曲状态系统 (7.27) 及对应的伴随系统.

引理 7.9(双曲线性 Neumann 问题的基本正则性) 令 $(\phi, u, y_0, y_1) \in L^1(0,T; L^2(\Omega)) \times L^2(\Sigma) \times H^1(\Omega) \times L^2(\Omega)$, 并设 $y(\phi, u, y_0, y_1) \in \mathcal{C}([0,T]; L^2(\Omega)) \cap \mathcal{C}^1([0,T]; H^1(\Omega)^*)$ 为线性 Neumann 边值问题 (7.31) 的唯一弱解. 则从 $L^2(\Sigma)$ 到 $\mathcal{C}([0,T]; H^{1/2}(\Omega)) \cap \mathcal{C}^1([0,T]; H^{1/2}(\Omega)^*)$ 的映射 $u \mapsto y(0,u,0,0)$ 是有界的, 它从 $L^2(\Sigma)$ 到 $H^{3/5-\varepsilon}(Q)$ 也是有界的 (对任意 $\varepsilon > 0$). 进一步, 从 $L^1(0,T; L^2(\Omega)) \times H^1(\Omega) \times L^2(\Omega)$ 到 $\mathcal{C}([0,T]; H^1(\Omega)) \cap \mathcal{C}^1([0,T]; L^2(\Omega))$ 的映射 $(\phi, y_0, y_1) \mapsto y(\phi, 0, y_0, y_1)$ 是有界的.

下面考虑线性波动方程的非齐次 Neumann 边值问题, 其数据可能是非光滑的:

$$\begin{cases} y_{tt} - \Delta y + \theta y = \phi, & \text{在 } Q \text{ 中}, \\ \partial_\nu y = u, & \text{在 } \Sigma \text{ 中}, \\ y(0) = y_0, \quad y_t(0) = y_1, & \text{在 } \Omega \text{ 中}, \end{cases} \tag{7.32}$$

其中非光滑系数 $\theta(x,t)$ 属于 $L^\infty(Q)$. 下述 (7.32) 中齐次线性 Neumann 边值问题的弱解的估计在这个系列中是需要的.

引理 7.10(齐次非光滑线性 Neumann 问题解的估计) 设 $u = 0$ 且初始数据 (ϕ, y_0, y_1) 属于 $L^1(0,T; L^2(\Omega)) \times H^1(\Omega) \times L^2(\Omega)$. 则 (7.32) 中的齐次 Neumann 问题在 $\mathcal{C}([0,T]; L^2(\Omega)) \cap \mathcal{C}^1([0,T]; H^1(\Omega))$ 中拥有唯一弱解. 该解满足估计

$$\begin{aligned} &\|y\|_{\mathcal{C}([0,T];H^1(\Omega))} + \|y_t\|_{\mathcal{C}([0,T];L^2(\Omega))} \\ &\leqslant C\big(\|\phi\|_{L^1(0,T;L^2(\Omega))} + \|y_0\|_{H^1(\Omega)} + \|y_1\|_{L^2(\Omega)} \end{aligned}$$

其中常数 $C > 0$ 可以依赖于 $\|\theta\|_{L^\infty(Q)}$ 和 $\|\phi\|_{L^1(0,T;L^2(\Omega))}$, 但对具有相同的 $L^\infty(Q)$-范数的 $\theta(x,t)$ 是不变的.

证明 证明是标准的, 只需要在 (7.32) 中的第一个方程乘以 y_t, 在 Ω 上积分, 然后应用经典的 Gronwall 引理; 比如, Lions 的书 [791] 可见更多细节. △

下面的引理建立了 (7.32) 中的非光滑非齐次线性 Neumann 问题的控制/弱解算子的一个重要的紧性性质.

引理 7.11(非光滑非齐次线性 Neumann 问题弱解的紧性) 设 $(\phi, y_0, y_1) = (0,0,0)$ 且 $u \in L^2(\Sigma)$. 则 (7.32) 中的非齐次 Neumann 问题具有唯一弱解 $y(u)$ 属于 $\mathcal{C}\big([0,T];L^2(\Omega)\big)\cap\mathcal{C}^1\big([0,T];H^1(\Omega)^*\big)$, 使得解映射 $u \mapsto \big(y(u), y_t(u)\big)$ 是从 $L^2(\Sigma)$ 到乘积空间 $\mathcal{C}\big([0,T];H^{1/2}(\Omega)\big)\times\mathcal{C}\big([0,T];H^{1/2}(\Omega)\big)$ 的有界算子. 进一步, 映射 $u \mapsto y(u)$ 是 $L^2(\Sigma)$ 到 $\mathcal{C}\big([0,T];L^2(\Omega)\big)$ 的紧算子.

证明 在 $L^2\big(0,\bar{t};L^2(\Omega)\big)$ 中当 $\bar{t}$ 足够小时应用不动点方法, 并重复这个过程 m 次 (m 满足 $m\bar{t} > T$), 则 (7.32) 弱解的存在性和唯一性可由线性系统 (7.31) 的结果得到. 由此得

$$\|y\|_{\mathcal{C}([0,T];H^{1/2}(\Omega))} + \|y_t\|_{\mathcal{C}([0,T];H^{1/2}(\Omega))} \leqslant C\|u\|_{L^2(\Sigma)},$$

其中 $C > 0$ 依赖于范数 $\|\theta\|_{L^\infty(Q)}$ 的一个上界, 但不依赖于函数 $\theta(\cdot)$ 本身. 紧性结果可由 Simon [1212, 推论 5] 直接得出. △

下面研究对原来的半线性波动方程的 Neumann 边值问题 (7.27), 方便起见称之为"状态系统."首先澄清 (7.27) 中非线性 Neumann 问题弱解这个术语的适合本节研究的定义.

定义 7.12(Neumann 状态系统的弱解) 满足 $(y, y_t)\in\mathcal{C}\big([0,T];L^2(\Omega)\big)\times\mathcal{C}\big([0,T];H^1(\Omega)^*\big)$ 的函数 $y(\cdot)$ 称为状态系统 (7.27) 的"弱解", 如果

$$\begin{aligned}&\int_Q -\vartheta(\cdot,y)z\,\mathrm{d}x\,\mathrm{d}t\\&=\int_Q y\varphi\,\mathrm{d}x\,\mathrm{d}t - \langle y_t(0), z(0)\rangle_{H^1(\Omega)^*,H^1(\Omega)} + \int_Q y(0)z_t(0)\,\mathrm{d}x + \int_\Sigma zu\,\mathrm{d}s\,\mathrm{d}t\end{aligned}$$

对任意 $\varphi \in L^1\big(0,T;L^2(\Omega)\big)$ 成立, 其中 $z(\cdot)$ 是如下齐次 Neumann 边值问题的解:

$$\begin{cases} z_{tt} - \Delta z = \varphi, & \text{在 } Q \text{ 中},\\ \partial_\nu z = 0, & \text{在 } \Sigma \text{ 中},\\ z(T) = 0, \quad z_t(T) = 0, & \text{在 } \Omega \text{ 中}.\end{cases}$$

上述定义的优点是, 在 7.2.1 小节的常驻假设下, 可以建立原来状态系统弱解的存在性、唯一性和正则性.

定理 7.13(Neumann 状态系统弱解的存在性、唯一性和正则性) 对任意初始三元组 $(u, y_0, y_1) \in L^2(\Sigma)\times H^1(\Omega)\times L^2(\Omega)$, 状态系统 (7.27) 具有满足 $(y, y_t) \in \mathcal{C}\big([0,T];L^2(\Omega)\big)\times\mathcal{C}\big([0,T];H^1(\Omega)^*\big)$ 的唯一弱解 $y(\cdot)$, 使得 (y, y_t) 属于 $\mathcal{C}\big([0,T];$

$H^{1/2}(\Omega)) \times \mathcal{C}([0,T]; H^{1/2}(\Omega)^*)$ 并满足估计

$$\begin{aligned}&\|y\|_{\mathcal{C}([0,T];H^{1/2}(\Omega))} + \|y_t\|_{\mathcal{C}([0,T];H^{1/2}(\Omega)^*)}\\ \leqslant &C\big(\|u\|_{L^2(\Sigma)} + \|y_0\|_{H^1(\Omega)} + \|y_1\|_{L^2(\Omega)} + 1\big)\end{aligned}$$

对某 $C>0$ 成立. 进一步, 从 $L^2(\Sigma)\times H^1(\Omega)\times L^2(\Omega)$ 映入 $\mathcal{C}([0,T];H^{1/2}(\Omega))\cap \mathcal{C}^1([0,T];H^{1/2}(\Omega)^*)$ 的映射 $(u,y_0,y_1)\mapsto y$ 是连续的.

证明 当 $\bar{t}$ 足够小时, 由标准的不动点方法可以得到状态系统在交空间 $\mathcal{C}([0,\bar{t}]; L^2(\Omega))\cap\mathcal{C}^1([0,\bar{t}];H^1(\Omega)^*)$ 中弱解的存在性. 由此及假设 (H1) 与引理 7.10 和引理 7.11 中的估计就可以导出在定理中所述泛函空间中解的存在性. 解的唯一性的证明也是标准的, 简单起见这里也略去. (y,y_t) 在 $\mathcal{C}([0,T];H^{1/2}(\Omega))\cap\mathcal{C}^1([0,T];H^{1/2}(\Omega)^*)$ 中的估计根据引理 7.9 的基本正则性由 y 在 $\mathcal{C}([0,T];L^2(\Omega))$ 中的估计得出. 最后, 再用一次假设 (H1) 以及引理 7.10 和引理 7.11 中线性化了的系统 (7.32) 的对应估计, 则得从 $L^2(\Sigma)\times H^1(\Omega)\times L^2(\Omega)$ 映入 $\mathcal{C}([0,T];H^{1/2}(\Omega))\cap\mathcal{C}^1([0,T];H^{1/2}(\Omega)^*)$ 的映射 $(u,y_0,y_1)\mapsto y$ 的连续性. △

接下来考虑 (7.27) 的 (线性化了的) 伴随系统

$$\begin{cases} p_{tt}-\Delta p+\theta p=\mu|_Q, & \text{在 } Q \text{ 中},\\ \partial_\nu p=0, & \text{在 } \Sigma \text{ 中},\\ p(T)=0,\quad p_t(T)=-\mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中}.\end{cases}\tag{7.33}$$

其中 $\mu\in\mathcal{M}_b(]0,T];L^2(\Omega))$, $\mu|_Q$ 和 $\mu|_{\Omega\times\{T\}}$ 表示 μ 在 Q 和 $\Omega\times\{T\}$ 上的限制, $\theta(x,t)\in L^\infty(Q)$ 如 (7.32). 为引入并阐释伴随系统 (7.33) 具有适定性质弱解的恰当定义, 需要下述引理, 它显然也有其独立的意义.

引理 7.14(散度公式) 装备了范数

$$\|V\|_W := \|V\|_{(L^2(Q))^{n+1}} + \|\mathrm{div}\,(V)\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}$$

的泛函空间

$$W := \Big\{V\in\big(L^2(Q)\big)^{n+1} \mid \mathrm{div}\,(V)\in\mathcal{M}_b(]0,T[;L^2(\Omega))\Big\}$$

是 Banach 空间. 进一步, 存在唯一从 W 到 $H^{-1/2}(\partial Q)$ 的连续算子 γ_{ν_Q} 满足

$$\gamma_{\nu_Q}(V)=\gamma_0(V)\cdot\nu_Q \quad (\text{当} V\in\big(\mathcal{C}^1(\overline{Q})\big)^{n+1}),$$

并且散度公式

$$\begin{aligned}&\int_Q V\cdot\nabla\phi + \Big\langle \phi, \mathrm{div}\,(V)\Big\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T[;L^2(\Omega))}\\ =&\Big\langle \gamma_{\nu_Q}(V),\gamma_0(\phi)\Big\rangle_{H^{-1/2}(\partial Q),H^{1/2}(\partial Q)}\end{aligned}$$

对任意 $\phi \in H^1(Q)$ 成立, 其中 ∂Q 按通常约定表示 Q 的边界.

证明 易证 W 在所给的范数下是 Banach 空间. 令 Λ 为一个从 $H^{1/2}(\partial Q)$ 到 $H^1(Q)$ 的线性扩张算子, 即其为从 $H^{1/2}(\partial Q)$ 到 $H^1(Q)$ 的有界算子且

$$\gamma_0 \Lambda\varphi = \varphi, \quad \forall\varphi \in H^{1/2}(\partial Q).$$

取 $V \in \left(\mathcal{C}^1(\overline{Q})\right)^{n+1}$, 注意到泛函

$$\varphi \mapsto \int_Q V \cdot \nabla\Lambda\varphi + \left\langle \Lambda\varphi, \operatorname{div}(V)\right\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T[;L^2(\Omega))}$$

在 $H^{1/2}(\partial Q)$ 上是线性有界的. 记此泛函为 $\gamma_{\nu_Q}(V)$, 则可直接验证

$$\gamma_{\nu_Q}(V) = \gamma_0(V) \cdot \nu_Q$$

且定理中的散度公式成立. 这说明 $\gamma_{\nu_Q}(V)$ 不依赖于扩张算子 Λ. 进一步有

$$\left| \int_Q V \cdot \nabla\Lambda\varphi + \langle \lambda\varphi, \operatorname{div}(V)\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T[;L^2(\Omega))} \right| \leqslant C\|\varphi\|_{H^{1/2}(\partial Q)}\|V\|_W,$$

这蕴涵估计

$$\|\gamma_{\nu_Q}(V)\|_{H^{-1/2}(\partial Q)} \leqslant C\|V\|_W, \quad \forall V \in \left(\mathcal{C}^1(\overline{Q})\right)^{n+1}.$$

因为 $\left(\mathcal{C}^1(\overline{Q})\right)^{n+1}$ 在 W 中稠密, 定理证毕. △

下面取 $(p, p_t) \in L^2(0,T;H^1(\Omega)) \times L^2(0,T;L^2(\Omega))$ 并假设在 Q 上的分布的意义下计算的组合 $p_{tt} - \Delta p$ 属于 $\mathcal{M}_b(]0,T[;L^2(\Omega))$. 由引理 7.14, 在 ∂Q 上定义向量场 $(-\nabla p, p_t)$ 的"法向迹"为 $H^{-1/2}(\partial Q)$ 中的元. 则有

$$\begin{aligned}
&\|\gamma_{\nu_Q}(-\nabla p, p_t)\|_{H^{-1/2}(\partial Q)} \\
\leqslant\ & C\left(\|p\|_{L^2(0,T;H^1(\Omega))} + \|p_t\|_{L^2(Q)} + \|p_{tt} - \Delta p\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}\right),
\end{aligned}$$

其中 $C > 0$ 不依赖于 p. 因为 $\Omega \times \{0\}$ 是 ∂Q 的开子集, 算子 $\gamma_{\nu_Q}(-\nabla p, p_t)$ 在 $\Omega \times \{0\}$ 上的限制属于空间 $H^{-1/2}(\Omega)$. 因此有

$$\gamma_{\nu_Q}(-\nabla p, p_t)|_{\Omega\times\{0\}} = p_t(0) \in H^{-1/2}(\Omega).$$

可以看到, 这个结果是可以改进的. 在定理 7.16 中将证明, 伴随系统 (7.33) 的一个恰当定义的解 $p(\cdot)$ 事实上具有性质 $p_t(0) \in L^2(\Omega)$.

现在可以引入伴随系统 (7.33) 一个恰当弱解的定义并阐释其性质.

定义 7.15(Neumann 伴随系统的弱解) 称函数 $p \in L^\infty(0,T;L^2(\Omega))$ 为 (7.33) 的"弱解", 如果

$$\langle y(\varphi), \mu\rangle_{\mathcal{C}([0,T];L^2(\Omega))\times\mathcal{M}_b(]0,T];L^2(\Omega))} - \int_Q p\varphi\,\mathrm{d}x\mathrm{d}t = 0 \tag{7.34}$$

对任意 $\varphi \in L^1\big(0,T;L^2(\Omega)\big)$ 成立, 其中 $y(\varphi)$ 是下述系统的解:

$$\begin{cases} y_{tt} - \Delta y + \vartheta y = \varphi, & \text{在 } Q \text{ 中}, \\ \partial_\nu y = 0, & \text{在 } \Sigma \text{ 中}, \\ y(0) = 0, \quad y_t(0) = 0, & \text{在 } \Omega \text{ 中}. \end{cases} \tag{7.35}$$

在所给的常驻假设下, 下述定理给出了伴随系统 (7.33) 弱解的存在性、唯一性和正则性. 注意 $\mathcal{C}_w\big([0,T];H^1(\Omega)\big)$ 表示装备了弱拓扑的从 $[0,T]$ 到 $H^1(\Omega)$ 的连续函数空间.

定理 7.16(Neumann 伴随系统弱解的存在性、唯一性和正则性) 在所给的常驻假设下, 伴随系统 (7.33) 具有唯一弱解 $p(\cdot)$ 满足 $(p,p_t) \in L^\infty\big(0,T;H^1(\Omega)\big) \times L^\infty\big(0,T;L^2(\Omega)\big)$,

$$p_t \in BV\big([0,T];H^1(\Omega)^*\big), \quad p \in \mathcal{C}_w\big([0,T];H^1(\Omega)\big),$$
$$p_t(\tau) \in L^2(\Omega), \quad \forall \tau \in \big\{t \in [0,T] \mid \mu(\{t\}) = 0\big\},$$

这蕴涵 $p_t(0) \in L^2(\Omega)$. 进一步有估计

$$\|p\|_{L^\infty(0,T;H^1(\Omega))} + \|p_t\|_{L^\infty(0,T;L^2(\Omega))} \leqslant C\|\mu\|_{\mathcal{M}_b(]0,T];L^2(\Omega))},$$

其中 C 依赖于 $\|\vartheta\|_{L^\infty(Q)}$, 但对 $L^\infty(Q)$ 中相同范数的 $\vartheta(x,t)$ 是不变的.

证明 注意到, 如果 $(p,p_t) \in L^\infty\big(0,T;H^1(\Omega)\big) \times L^\infty\big(0,T;L^2(\Omega)\big)$ 满足 (7.34)(其中 $\mu = 0$), 则 $p = 0$. 这说明伴随系统 (7.33) 至多有一个弱解. 为证明弱解的存在性, 下面建立一个逼近过程. 首先构造序列 $\{\mu_k\} \subset L^1\big(0,T;L^2(\Omega)\big)$ 满足

$$\begin{cases} \|\mu_k\|_{L^1(0,T;L^2(\Omega))} = \|\mu\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}, \\ \displaystyle\lim_{k\to\infty} \int_Q y\mu_k \,\mathrm{d}x\,\mathrm{d}t = \Big\langle y, \mu|_{]0,T[} \Big\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T[;L^2(\Omega))}, \\ \forall y \in \mathcal{C}\big([0,T];L^2(\Omega)\big). \end{cases}$$

现在述构造定义 μ_k. 令 $\bar{\mu}$ 为 $\mu|_{]0,T[}$ 在 $\mathbb{R}$ 上的扩张 ($]0,T[$ 外部分定义为零), $\{\rho_k\}$ 为 $\mathbb{R}$ 上非负对称光滑化算子序列, 其支集含于 $(-1/k,1/k)$, 并令 ψ_0 和 ψ_T 为 $\mathbb{R}$ 上的函数, 定义为 $\psi_0(t) := -t$ 和 $\psi_T(t) := 2T - t$. 给定 $k \geqslant 2$, 对每个 Borel 子集 $S \subset \mathbb{R}$ 置

$$\bar{\mu}_k(A) := \big(\bar{\mu} * \rho_k\big)(S) + \big(\bar{\mu} * \rho_k\big)\big(\psi_0(S)\big) + \big(\bar{\mu} * \rho_k\big)\big(\psi_T(S)\big),$$

其中 $*$ 表示 $\bar{\mu}$ 和正则化核 ρ_k 的卷积. 因为这两个分布都具有紧支集, 上述卷积是有意义的. 接着构造所需的测度为

$$\mu_k := \frac{\|\mu\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}}{\|\bar{\mu}_k|_{]0,T[}\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}} \bar{\mu}_k|_{]0,T[}.$$

可以验证, 序列 $\{\mu_k\} \subset L^1\big(0,T;L^2(\Omega)\big)$ 满足上面列出的两个关系.

现在考虑系统

$$\begin{cases} p_{tt} - \Delta p + \vartheta p = \mu_k, & \text{在 } Q \text{ 中}, \\ \partial_\nu p = 0, & \text{在 } \Sigma \text{ 中}, \\ p(T) = 0, \quad p_t(T) = -\mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中} \end{cases} \tag{7.36}$$

的唯一解 p_k 并应用引理 7.10, 则得估计

$$\begin{aligned} &\|p_k\|_{L^\infty(0,T;H^1(\Omega))} + \|p_{kt}\|_{L^\infty(0,T;L^2(\Omega))} + \|p_k(0)\|_{H^1(\Omega)} \\ &+ \|p_{kt}(0)\|_{L^2(\Omega)} \leqslant C\|\mu\|_{\mathcal{M}_b(]0,T];L^2(\Omega))}, \end{aligned} \tag{7.37}$$

其中 $C>0$ 为不依赖于 k 的常数, p_{kt} 表示 p_k 对于 $t\in(0,T)$ 的在向量值分布意义下的导数. 以 p_{ktt} 表示 p_{kt} 对于 $t\in(0,T)$ 的对应导数并应用 (7.36), 则有

$$\begin{aligned} p_{ktt} = \pi_k + \mu_k &\in L^\infty\big(0,t;H^1(\Omega)^*\big) + \mathcal{M}_b\big(]0,T[;L^2(\Omega)\big) \\ &\subset \mathcal{M}_b\big(]0,T[;H^1(\Omega)^*\big), \end{aligned}$$

其中算子 π_k 定义为

$$\langle \pi_k, y\rangle_{L^\infty(0,T;H^1(\Omega)^*),L^1(0,T;H^1(\Omega))} := \int_Q \big(\nabla p_k\cdot\nabla y - \vartheta p_k y\big)\,\mathrm{d}x\,\mathrm{d}t.$$

因此, 除了 (7.37) 成立, 序列 $\{p_{ktt}\}$ 和 $\{p_{kt}\}$ 还在空间 $\mathcal{M}_b\big(]0,T[;H^1(\Omega)^*\big)$ 和 $BV\big([0,T];H^1(\Omega)^*\big)$ 中分别是有界的. 注意到 $\mathcal{M}_b\big(]0,T[;H^1(\Omega)^*\big)$ 是一个可分 Banach 空间的对偶, 则可以在上述序列中选取弱 * 收敛的子序列. 同样的弱 * 列紧性质对 $BV\big([0,T];H^1(\Omega)^*\big)$ 也成立. 这样可找到 $p\in L^\infty\big(0,T;H^1(\Omega)\big)$ 满足 $p_t\in L^\infty\big(0,T;L^2(\Omega)\big)\cap BV\big([0,T];H^1(\Omega)^*\big)$ 和一个子序列 $\{p_k\}$ 在 $L^\infty\big(0,T;H^1(\Omega)\big)$ 中弱 * 收敛到 p, $\{p_{kt}\}$ 在 $L^\infty\big(0,T;L^2(\Omega)\big)$ 中弱 * 收敛到 p_t. 进一步, 因为 $\gamma_{\nu_Q}(-\nabla p_k, p_{kt})$ 在 $L^2(\partial Q)$ 中有界, 也可得 $\gamma_{\nu_Q}(-\nabla p_k, p_{kt})$ 在 $L^2(\partial Q)$ 中沿弱拓扑收敛到 $\gamma_{\nu_Q}(-\nabla p, p_t)$. 考虑关系

$$\gamma_{\nu_Q}(-\nabla p_k, p_{kt})|_{\Omega\times\{T\}} = \mu|_{\Omega\times\{T\}}, \quad \gamma_{\nu_Q}(-\nabla p_k, p_{kt})|_\Sigma = 0,$$

就得到 $\gamma_{\nu_Q}(-\nabla p, p_t)|_\Sigma = -\partial_\nu p = 0$ 和在 $L^2(\Omega)$ 的弱拓扑下有

$$\gamma_{\nu_Q}(-\nabla p_k, p_{kt})|_{\Omega\times\{0\}} = p_{kt}(0) \xrightarrow{w} \gamma_{\nu_Q}(-\nabla p, p_t)|_{\Omega\times\{0\}} = p_t(0).$$

最后, 在等式

$$\langle y(\varphi), \mu_k\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T];L^2(\Omega))} - \int_Q p_k\varphi\,\mathrm{d}x\,\mathrm{d}t = 0$$

取极限, 其中 $y(\varphi)$ 是 (7.35) 的解, 则得 $p(\cdot)$ 是所需的 (7.33) 的弱解. 证毕. △

在本小节的最后, 给出所考虑 Neumann 问题的 (线性化了的) 状态和伴随系统对应解之间的 Green 类型的关系.

定理 7.17(双曲 Neumann 问题的 Green 公式)　给定 $(\phi, y_0, y_1) = (0,0,0)$ 和 $u \in L^2(\Sigma)$, 考虑系统 (7.32) 的对应弱解 $y(\cdot)$. 给定 $\mu \in \mathcal{M}_b(]0,T]; L^2(\Omega))$, 设 p 满足伴随系统 (7.33). 则有

$$\langle y, \mu \rangle_{\mathcal{C}([0,T];L^2(\Omega)), \mathcal{M}_b(]0,T];L^2(\Omega))} - \int_Q p\varphi \,\mathrm{d}x\,\mathrm{d}t = \int_\Sigma pu \,\mathrm{d}s\,\mathrm{d}t.$$

证明　由定理 7.16 的证明可知, Green 公式的一个近似版本对二元组 (y, p_k) 成立, 其中 p_k 是逼近伴随系统 (7.36) 的对应弱解 $(k \in \mathbb{N})$. 在 $k \to \infty$ 时取极限, 则得定理中的 Green 公式. △

7.2.3　针形变分和增量公式

如前所述, 这里推导原状态约束 Neumann 问题 (NP) 必要最优条件的方法包括一个逼近过程来对状态约束应用罚函数. 该过程得到这样一类双曲方程的 Neumann 边界控制问题, 其控制具有逐点/硬的约束, 但没有状态约束. 尽管这些逼近得到的问题比原来的状态约束问题 (NP) 容易得多, 但还是需要一个细致的变分分析. 如在常微分最优控制的情形, 众所周知, 对控制具有硬约束而没有状态变量的问题, 推导最大值原理类型条件的关键是对针形变分来极小化目标函数的"增量公式"; 比照 6.3 节. 本节对所论的双曲控制问题建立一些类似结果, 这里使用了针形变分的高维版本, 在偏微分方程控制理论中一般称为参考控制的"扩散扰动"或"(多) 刺/补丁扰动". 本书使用"扩散"这个术语.

给定参考控制 $\bar{u}(\cdot) \in U_{ad}$, 容许控制 $u(\cdot) \in U_{ad}$, 及常数 $\rho \in (0,1)$, $\bar{u}$ 的一个"扩散扰动/变分"定义为

$$u_\rho(s,t) := \begin{cases} \bar{u}(s,t), & \text{在 } \Sigma \backslash E_\rho \text{ 中}, \\ u(s,t), & \text{在 } E_\rho \text{ 中}, \end{cases} \tag{7.38}$$

其中 E_ρ 是 Σ 的一个可测子集. 下述定理可理解为费用函数 $J(y,u)$ 相对于参考控制的扩散扰动的增量公式. 注意它还包含 (7.27) 的状态轨道的对应 Taylor 展开, 这是增量公式的一个要素. 下面用 $\widehat{\Delta} J$ 来表示费用函数 J 的增量以区别于 Laplace 算子 Δ.

定理 7.18(Neumann 问题的增量公式)　给定任意控制 $\bar{u}, u \in U_{ad}$ 和一个常数 $\rho \in (0,1)$, 考虑扩散扰动 (7.38) 和系统 (7.27) 分别对应于 $\bar{u}$ 和 u_ρ 的弱解 $\bar{y}$ 和 y_ρ. 则存在一个可测子集 $E_\rho \subset \Sigma$ 满足:

$$\mathcal{L}^n(E_\rho) = \rho \mathcal{L}^n(\Sigma),$$

$$\int_{E_\rho} \big(h(s,t,\bar{u}) - h(s,t,u)\big)\,\mathrm{d}s\,\mathrm{d}t = \rho \int_{\Sigma} \big(h(s,t,\bar{u}) - h(s,t,u)\big)\,\mathrm{d}s\,\mathrm{d}t,$$

$$y_\rho = \bar{y} + \rho z + \rho r_\rho, \qquad \lim_{\rho\to 0} \|r_\rho\|_{C([0,T];L^2(\Omega))} = 0, \tag{7.39}$$

$$\begin{aligned} &J(y_\rho, u_\rho) = J(\bar{y},\bar{u}) + \rho\widehat{\Delta} J + o(\rho), \\ &\text{其中} \quad \widehat{\Delta} J := J'_y(\bar{y},\bar{u})z + J(\bar{y},u) - J(\bar{y},\bar{u}), \end{aligned} \tag{7.40}$$

而 $z(\cdot)$ 是下面系统的弱解:

$$\begin{cases} z_{tt} - \Delta z + \vartheta'_y(\cdot,\bar{y})z = 0, & \text{在 } Q \text{ 中}, \\ \partial_\nu z = \bar{u} - u, & \text{在 } \Sigma \text{ 中}, \\ z(0) = 0, \quad z_t(0) = 0, & \text{在 } \Omega \text{ 中}. \end{cases} \tag{7.41}$$

下面给出的证明依赖于 Raymond 和 Zidani 在文章 [1121] 中建立的技术性引理 (文中引理 4.2), 读者可以在该文中找到所有细节. 注意这里 χ_E 指集合 E 的指示函数, 即其在 E 上取值为 1, 在 E 外取值为 0.

引理 7.19(扩散扰动的性质) 令 $\bar{u}, u \in U_{ad}$. 则对任意 $\rho \in (0,1)$, 存在可测集序列 $E^k_\rho \subset \Sigma$ 使得

$$\mathcal{L}^n(E^k_\rho) = \rho\mathcal{L}^n(\Sigma),$$

$$\int_{E^k_\rho} \big(h(s,t,\bar{u}) - h(s,t,u)\big)\,\mathrm{d}s\,\mathrm{d}t = \rho \int_{\Sigma} \big(h(s,t,\bar{u}) - h(s,t,u)\big)\,\mathrm{d}s\,\mathrm{d}t,$$

$$\frac{1}{\rho}\chi_{E^k_\rho} \xrightarrow{w^*} 1 \quad (\text{在 } L^\infty(\Sigma) \text{ 中, 其中} k \to \infty).$$

定理 7.18 的证明 满足定理要求的子集 E_ρ 的存在性是引理 7.19 的简单推论. 这里主要是证明 Taylor 展开 (7.39). 根据扩散扰动的构造, 这显然蕴涵增量公式 (7.40).

为证 (7.39), 取 $\rho \in (0,1)$, 从引理 7.19 找到集合 E^k_ρ, 并构造扩散控制扰动为

$$u^k_\rho(s,t) := \begin{cases} \bar{u}(s,t), & \text{在 } \Sigma\backslash E^k_\rho \text{ 中}, \\ u(s,t), & \text{在 } E^k_\rho \text{ 中}. \end{cases}$$

令 y^k_ρ 为 (7.27) 对应于 u^k_ρ 的解, 并令 z 为 (7.41) 的 (唯一) 弱解. 易知对任一 $\rho \in (0,1)$ 和 $k \in \mathbb{N}$, 函数 $\xi^k_\rho := \dfrac{1}{\rho}(y^k_\rho - \bar{y}) - z$ 是下面系统的唯一弱解:

$$\begin{cases} \xi_{tt} - \Delta\xi + \theta^k_\rho\xi = f^k_\rho, & \text{在 } Q \text{ 中}, \\ \partial_\nu \xi = w^k_\rho, & \text{在 } \Sigma \text{ 中}, \\ \xi(0) = 0, \quad \xi_t(0) = 0, & \text{在 } \Omega \text{ 中}. \end{cases}$$

其中的数据给出如下: $f_\rho^k := \big(\vartheta_y'(\cdot,\bar{y}) - \theta_\rho^k\big)z$,

$$\theta_\rho^k := \int_0^1 \vartheta_y'\big(\cdot,\bar{y}+\tau(y_\rho^k-\bar{y})\big)\,\mathrm{d}\tau, \quad w_\rho^k := \left(1-\frac{1}{\rho}\chi_{E_\rho^k}\right)(u-\bar{u}).$$

以 $\xi_\rho^{k,1}$ 记

$$\begin{cases} \xi_{tt}-\Delta\xi+\theta_\rho^k\xi = f_\rho^k, & \text{在 } Q \text{ 中},\\ \partial_\nu\xi = 0, & \text{在 } \varSigma \text{ 中},\\ \xi(0)=0, \quad \xi_t(0)=0, & \text{在 } \varOmega \text{ 中} \end{cases}$$

的解, 以 $\xi_\rho^{nk,2}$ 记

$$\begin{cases} \xi_{tt}-\Delta\xi+\theta_\rho^k\xi = 0, & \text{在 } Q \text{ 中},\\ \partial_\nu\xi = w_\rho^k, & \text{在 } \varSigma \text{ 中},\\ \xi(0)=0, \quad \xi_t(0)=0, & \text{在 } \varOmega \text{ 中} \end{cases}$$

的解, 并以 ζ_ρ^k 记

$$\begin{cases} \zeta_{tt}-\Delta\zeta+\theta\zeta = 0, & \text{在 } Q \text{ 中},\\ \partial_\nu\zeta = w_\rho^k, & \text{在 } \varSigma \text{ 中},\\ \zeta(0)=0, \quad \zeta_t(0)=0, & \text{在 } \varOmega \text{ 中} \end{cases}$$

的解, 其中 $\theta(x,t) := \vartheta_y'\big(x,t,\bar{y}(x,t)\big)$. 则显然有

$$\begin{aligned} &(\xi_\rho^{k,2}-\zeta_\rho^k)_{tt} - \Delta\big(\xi_\rho^{k,2}-\zeta_\rho^k\big) + \theta_\rho^k\big(\xi_\rho^{k,2}-\zeta_\rho^k\big) = \big(\theta-\theta_\rho^k\big)\zeta_\rho^k, && \text{在 } Q \text{ 中},\\ &\partial_\nu\big(\xi_\rho^{k,2}-\zeta_\rho^k\big) = 0, && \text{在 } \varSigma \text{ 中},\\ &\big(\xi_\rho^{nk,2}-\zeta_p^k\big)(0) = 0, \quad \big(\xi_\rho^{k,2}-\zeta_\rho^k\big)_t(0) = 0, && \text{在 } \varOmega \text{ 中}. \end{aligned}$$

由引理 7.10, 可找到与 k 和 ρ 无关的常数 $C>0$ 使得下述估计

$$\begin{aligned} \big\|\xi_\rho^{k,2}-\zeta_\rho^k\big\|_{\mathcal{C}([0,T];L^2(\varOmega))} &\leqslant C\big\|\theta-\theta_\rho^k\big\|_{L^1(0,T;L^{2n}(\varOmega))}\cdot\big\|\zeta_\rho^k\big\|_{L^\infty(0,T;L^{2n/(n-1)}(\varOmega))}\\ &\leqslant C\big\|\theta-\theta_\rho^k\big\|_{L^1(0,T;L^{2n}(\varOmega))}\cdot\big\|\zeta_\rho^k\big\|_{L^\infty(0,T;H^{1/2}(\varOmega))},\\ \big\|\xi_\rho^{k,1}\big\|_{C([0,T];L^2(\varOmega))} &\leqslant C\big\|f_\rho^k\big\|_{L^1(0,T;L^2(\varOmega))} \end{aligned}$$

对任意 $k\in\mathbb{N}$ 和 $0<\rho<1$ 成立, 其中根据引理 7.9, 函数 $\big\|\zeta_\rho^{nk}\big\|_{L^\infty(0,T;L^{2n/(n-1)}(\varOmega))}$ 一致有界. 由引理 7.19 中的弱 * 收敛性, w_ρ^k 在 $L^2(\varSigma)$ 中的弱拓扑下收敛到 0, 且由引理 7.11 得 ζ_ρ^k 在 $\mathcal{C}\big([0,T];L^2(\varOmega)\big)$ 中强收敛到 0 ($k\to\infty$, $\forall 0<\rho<1$). 因此存在整数 $k(\rho)$ 满足

$$\big\|\zeta_\rho^{k(\rho)}\big\|_{\mathcal{C}([0,T];L^2(\varOmega))} \leqslant \rho \quad (\text{当} 0<\rho<1).$$

进一步观察到 $u_\rho^{k(\rho)}$ 在 $L^2(\varSigma)$ 中强收敛到 $\bar{u}$ ($\rho\downarrow 0$). 所以从引理 7.13 知 $y_\rho^{k(\rho)}$ 在 $\mathcal{C}\big([0,T];L^2(\varOmega)\big)$ 中强收敛到 $\bar{y}$ ($\rho\downarrow 0$). 应用假设 (H1), 则得函数 $f_\rho^{k(\rho)}$ 在

$L^1(0,T;L^2(\Omega))$ 中强收敛到零, 而函数 $(\theta-\theta_\rho^{k(\rho)})$ 在 $L^1(0,T;L^{2n}(\Omega))$ 中强收敛到零 $(\rho\downarrow 0)$. 由上述估计, 这蕴涵关系

$$\lim_{\rho\to 0}\|\xi_\rho^{k(\rho)}\|_{\mathcal{C}([0,T];L^2(\Omega))}\leqslant\lim_{\rho\to 0}\Big(\|\xi_\rho^{k(\rho),1}\|_{\mathcal{C}([0,T];L^2(\Omega))}+\|\xi_\rho^{k(\rho),2}-\zeta_\rho^{k(\rho)}\|_{\mathcal{C}([0,T];L^2(\Omega))}+\|\zeta_\rho^{k(\rho)}\|_{\mathcal{C}([0,T];L^2(\Omega))}\Big)=0.$$

最后置 $E_\rho:=E_\rho^{nk(\rho)}$, $u_\rho:=u_\rho^{k(\rho)}$, 及 $\dfrac{1}{\rho}r_\rho:=\xi_\rho^{k(\rho)}$, 则完成证明. △

7.2.4 必要最优条件的证明

本节证明定理 7.8 给出的状态约束 Neumann 边界控制问题 (NP) 的必要最优条件. 证明涉及一个强逼近过程来构造状态约束的罚函数, 基于定理 2.26 给出的 Ekeland 变分原理. 为此首先描述一个完备度量空间和一个下半连续函数, 它们适合 Ekeland 变分原理的应用.

给定 $\bar{u}(\cdot)\in U_{ad}$ 和固定正常数 m, 定义集合

$$U_{ad}(\bar{u},m):=\{u\in U_{ad}\mid |u(s,t)-\bar{u}(s,t)|\leqslant m,\quad \text{a.e. }(s,t)\in\Sigma\}$$

并装备距离 $d(\cdot,\cdot)$ 如下:

$$d(v,u):=\mathcal{L}^N(\{(s,t)\mid v(s,t)\neq u(s,t)\}),$$

其中 $\mathcal{L}^n(\Omega)$ 和以前一样表示集合 $\Omega\subset\mathbb{R}^n$ 上的 n 维 Lebesgue 测度. 注意到如果 $\{u_k\}\subset U_{ad}(\bar{u},m)$ 和 $u\in U_{ad}(\bar{u},m)$ 满足 $\lim\limits_{k\to\infty}d(u_k,u)=0$, 则序列 $\{u_k\}$ 沿 $L^2(\Sigma)$ 的范数强收敛到 u. 下面的结果给出了这个空间及其上的 (NP) 的费用函数的更多信息, 其中 y_u 表示 (7.27) 对应于 u 的弱解.

引理 7.20(Eleland 原理所需的恰当配置) 度量空间 $(U_{ad}(\bar{u},m),d)$ 是完备的, 映射 $u\mapsto(y_u,J(y_u,u))$ 从 $(U_{ad}(\bar{u},m),d)$ 到 $\mathcal{C}([0,T];L^2(\Omega))\times\mathbb{R}$ 是连续的.

证明 空间 $(U_{ad}(\bar{u},m),d)$ 的完备性是众所周知的事实, 它可以追溯到 Ekeland 的原始文章 [397]. 现在基于 7.2.2 小节建立的状态系统 (7.27) 弱解的正则性来证明引理中的连续性结果.

为此取 $\{u_k\}\subset U_{ad}(\bar{u},m)$ 和 $u\in U_{ad}(\bar{u},m)$ 使得 $\{u_k\}$ 沿上述 d-度量在 $k\to\infty$ 时收敛到 u. 以 y 和 y_k 分别记 (7.27) 对应于 u 和 u_k 的弱解. 因为 $u_k\to u$ ($L^2(\Sigma)$ 中强收敛), 根据定理 7.13, 轨道 y_k 在 $\mathcal{C}([0,T];L^2(\Omega))$ 中强收敛到 y. 进一步根据假设 (H2)~(H4) 中的估计可得, 值序列 $J(y_k,u_k)$ 收敛到 $J(y,u)$ $(k\to\infty)$. 这就证明了连续性. △

现在应用 Banach 空间几何的一些经典结果 (例如可见 Li 和 Yong 的书 [789], 第二章), 由 $\mathcal{C}([0,T];L^2(\Omega))$ 的可分性可得该空间上存在等价范数 $|\cdot|_{\mathcal{C}([0,T];L^2(\Omega))}$, 该范

数在非零点Gâteaux可微, 且在 $\mathcal{M}([0,T];L^2(\Omega))$ 上的对偶范数 (记为 $|\cdot|_{\mathcal{M}([0,T];L^2(\Omega))}$) 为严格凸. 给定原问题 (NP) 的约束集 $\Theta\subset\mathcal{C}([0,T];L^2(\Omega))$, 考虑对应的用 $\mathcal{C}([0,T];L^2(\Omega))$ 上的新范数 $|\cdot|_{\mathcal{C}([0,T];L^2(\Omega))}$ 定义的距离函数

$$d_\Theta(x):=\inf_{z\in\Theta}|x-z|_{\mathcal{C}([0,T];L^2(\Omega))}.$$

则其为具有模 $\ell=1$ 的全局 Lipschitz 函数, 并由 Θ 的凸性知其在 $\mathcal{C}([0,T];L^2(\Omega))$ 上是凸的. 进一步有

$$\begin{cases}|\xi|_{\mathcal{M}([0,T];L^2(\Omega))}\leqslant 1 & (\text{若 } x^*\in\partial d_\Theta(x),\quad x\in\Theta),\\ |x^*|_{\mathcal{M}([0,T];L^2(\Omega))}=1 & (\text{若 } x^*\in\partial d_\Theta(x),\quad x\notin\Theta);\end{cases}$$

请比照 1.3.3 小节. 注意到对偶范数 $|\cdot|_{\mathcal{M}([0,T];L^2(\Omega))}$ 也是严格凸的, 则得次微分 $\partial d_\Theta(x)$ 是一个单点集, 从而 d_Θ 在 x 点 Gâteaux可微对任意 $x\notin\Theta$ 成立.

令 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$ 为原问题 (NP) 的一个最优解. 由上述距离函数 d_Θ 定义罚函数

$$J_m(y,u):=\left[\left(J(y,u)-J(\bar{y},\bar{u})+\frac{1}{m^2}\right)^+\right]^2+d_\Theta^2(y),\quad m\in\mathbb{N},$$

其中 J 是 (NP) 中的费用函数. 因为 $J_m(\bar{y},\bar{u})=m^{-4}$, 所以

$$J_m(\bar{y},\bar{u})<\inf\Big\{J_m(y,u)\mid u\in U_{ad}(\bar{u},m^{1/3}),\ (y,u)\text{ 满足}(7.27)\Big\}+\frac{1}{m^2}$$

对任意 $m\in\mathbb{N}$ 成立, 即 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$ 是罚问题的近似 $\dfrac{1}{m^2}$-最优解.

注意到 J_m 在其值非零的地方光滑 (在 Gâteaux可微的意义下); 比照定理 2.8 和定理 2.10 中极点原理证明里度量逼近中的光滑化过程. 这源于 J_m 的构造, 假设 (H2)~(H4), 以及上述 d_Θ 的性质. 对参考二元组 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$, 由 Ekeland 原理可找到强逼近 $\{y_m(\cdot),u_m(\cdot)\}$ 满足 (7.27), 且使 $\{y_m(\cdot),u_m(\cdot)\}$ 为 (7.27) 某扰动最优控制问题的确切解, 其中控制约束相同, 没有状态约束. 现在可以证明主要定理如下.

定理 7.8 的证明 证明如下分成三个主要步骤.

第一步: 以 Ekeland 原理逼近原问题. 给定原问题 (NP) 的最优解 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$, 固定自然数 $m\in\mathbb{N}$ 并由引理 7.20 知, $(U_{ad}(\bar{u},m^{1/3}),d)$ 为完备度量空间, 且函数 $u\mapsto J_m(y_u,u)$ 在该空间上下半连续 (其实是连续). 根据 Ekeland 变分原理, 可找到容许控制 u_m 满足

$$\begin{aligned}&u_m\in U_{ad}(\bar{u},m^{1/3}),\quad d(u_m,\bar{u})\leqslant\frac{1}{m},\\&J_m(y_m,u_m)\leqslant J_m(y_u,u)+\frac{1}{m}d(u_m,u),\end{aligned}\tag{7.42}$$

其中 $u \in U_{ad}(\bar{u}, m^{1/3})$ 任意, y_m 和 y_u 分别是 (7.27) 对应于 u_m 和 u 的弱解. 这意味着, 对任意 $m \in \mathbb{N}$, 控制 u_m 是下面扰动问题 (NP_m) 的最优解:

$$\inf \left\{ J_m(y,u) + \frac{1}{m} \mid u \in U_{ad}(\bar{u}, m^{1/3}),\ (y,u) \text{ 满足(7.27)} \right\}.$$

第二步: 逼近问题的必要条件. 首先任取控制 $u_0 \in U_{ad}$ 并构造 (NP) 的最优控制 $\bar{u}$ 的下述修改:

$$u_{0m}(s,t) := \begin{cases} u_0(s,t), & \text{若 } |u_0(s,t) - \bar{u}(s,t)| \leqslant m^{1/3}, \\ \bar{u}(s,t), & \text{其他}. \end{cases}$$

注意到 u_{0m} 是逼近问题 (NP_m) 的可行解 ($m \in \mathbb{N}$). 给定 $0 < \rho < 1$, 则可定义 (NP_m) 最优控制 u_m 的扩散扰动

$$u_\rho^m(s,t) := \begin{cases} u_m(s,t), & \text{在 } \Sigma \backslash E_\rho^m \text{ 中}, \\ u_{0m}(s,t), & \text{在 } E_\rho^m \text{ 中}. \end{cases}$$

定理 7.18 保证了可测集 $E_\rho^m \subset \Sigma$ 的存在性, 使得

$$\begin{aligned} &\mathcal{L}^n(E_\rho^m) = \rho \mathcal{L}^n(\Sigma), \quad y_\rho^m = y_m + \rho z_m + \rho r_\rho^m, \\ &\lim_{\rho \to 0} \left\| r_\rho^m \right\|_{C([0,T];L^2(\Omega))} = 0, \\ &J(y_\rho^m, u_\rho^m) = J(y_m, u_m) + \rho \widehat{\Delta} J^m + o(\rho), \end{aligned} \tag{7.43}$$

其中 y_ρ^m 是 (7.27) 对应于 u_ρ^m 的弱解, z_m 是系统

$$\begin{cases} z_{tt} - \Delta z + \vartheta_y'(\cdot, y_m) z = 0, & \text{在 } Q \text{ 中}, \\ \partial_\nu z = u_m - u_{0m}, & \text{在 } \Sigma \text{ 中}, \\ z(0) = 0, \quad z_t(0) = 0, & \text{在 } \Omega \text{ 中} \end{cases}$$

的弱解, 而 $\widehat{\Delta} J^m$ 定义为

$$\begin{aligned} \widehat{\Delta} J^m := &\int_Q g_y'(\cdot, y_m) z_m \,\mathrm{d}x\,\mathrm{d}t + \int_\Omega f_y'(\cdot, y_m(T)) z_m \,\mathrm{d}x \\ &+ \int_\Sigma \big(h(\cdot, u_{0m}) - h(\cdot, u_m)\big) \,\mathrm{d}s\,\mathrm{d}t. \end{aligned}$$

因为每个 u_ρ^m 对 (NP_m) 都是可行的, 由 (7.42) 和其中的度量 $d(\cdot,\cdot)$ 的构造得

$$\lim_{\rho \to 0} \frac{J_m(y_m, u_m) - J_m(y_\rho^m, u_\rho^m)}{\rho} \leqslant \frac{1}{m} \mathcal{L}^n(\Sigma). \tag{7.44}$$

因为 u_m 是 (NP_m) 的最优解, 并由 J_m 的构造知 $J_m(y_m, u_m) \neq 0$ ($m \in \mathbb{N}$). 所以由上述知 J_m 在 (y_m, u_m) 点 Gâteaux可微. 因此由 (7.43) 和 (7.44) 易得最优条件

$$-\lambda_m \widehat{\Delta} J^m - \langle \mu_m, z_m \rangle \leqslant \frac{1}{m} \mathcal{L}^n(\Sigma), \tag{7.45}$$

其中乘子 λ_m 和 μ_m 如下计算:

$$\lambda_m := \frac{\left(J(y_m, u_m) - J(\bar{y}, \bar{u}) + \dfrac{1}{m^2}\right)^+}{J_m(y_m, u_m)},$$

$$\mu_m := \begin{cases} \dfrac{d_\Theta(y_m)\nabla d_\Theta(y_m)}{J_m(y_m, u_m)}, & \text{若 } y_m \notin \Theta, \\ 0, & \text{其他}. \end{cases}$$

注意到 $\mu_m \in \mathcal{M}\big([0,T]; L^2(\Omega)\big)$, 考虑伴随系统

$$\begin{cases} p_{tt} - \Delta p + \vartheta'_y(\cdot, y_m)p = \lambda_m g'_y(\cdot, y_m) + \mu_m|_Q, & \text{在 } Q \text{ 中}, \\ \partial_\nu p = 0, & \text{在 } \Sigma \text{ 中}, \\ p(T) = 0, \quad p_t(T) = -\lambda_m f'_y\big(\cdot, y_m(T)\big) - \mu_m|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中} \end{cases}$$

的 (唯一) 弱解 p_m, 其中 $\mu_m|_Q$ 和 $\mu_m|_{\Omega\times\{T\}}$ 分别是 μ_m 在 Q 和 $\Omega\times\{T\}$ 上的限制. 由定理 7.17 中的 Green 公式得

$$\begin{aligned} &\lambda_m \int_Q g'_y(x,t,y_m) z_m \,\mathrm{d}x\,\mathrm{d}t + \lambda \int_\Omega f'_y\big(x, y_m(T)\big) z_m(T)\,\mathrm{d}x + \langle \mu_m, z_m\rangle \\ =& \int_Q p_m\big(z_{ktt} - \Delta z_m + \vartheta'_y(\cdot, y_m) z_m\big)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma p_m \partial_\nu z_m \,\mathrm{d}s\,\mathrm{d}t \\ =& \int_\Sigma p_m(u_m - u_{0m})\,\mathrm{d}s\,\mathrm{d}t. \end{aligned}$$

根据 (7.45) 和 $\widehat{\Delta} J^m$ 的定义, 最后一式蕴涵

$$\begin{aligned} &\int_\Sigma \big(\lambda_m h(s,t,u_m) + p_m u_m\big)\,\mathrm{d}s\,\mathrm{d}t \\ \leqslant& \int_\Sigma \big(\lambda_m h(s,t,u_{0m}) + p_m u_{0m}\big)\,\mathrm{d}s\,\mathrm{d}t + \frac{1}{m}\mathcal{L}^n(\Sigma), \quad \forall m \in \mathbb{N}, \end{aligned} \tag{7.46}$$

这就给出了欲证的逼近问题 (NP_m) 最优解 u_m 的必要最优条件.

第三步: 取极限.　最后需要对 (NP_m) 的最优解 u_m 的上述关系在 $m \to \infty$ 时取极限. 首先注意到

$$\lambda_m^2 + |\mu_m|^2_{\mathcal{M}([0,T];L^2(\Omega))} = 1, \quad \forall m \in \mathbb{N}.$$

由基本的泛函分析可得 $(\lambda, \bar{\mu}) \in \mathbb{R}\times\mathcal{M}\big([0,T]; L^2(\Omega)\big)$ (其中 $\lambda \geqslant 0$) 和序列 (λ_m, μ_m) 的子序列 (仍然用 m 作为下标) 满足

$$\lambda_m \to \lambda \quad (\text{在 } \mathbb{R} \text{ 中}), \quad \mu_m \xrightarrow{w^*} \bar{\mu} \quad (\text{即沿}\mathcal{M}([0,T]; L^2(\Omega))\text{的弱 * 拓扑}).$$

进一步, 定理 7.16 保证了估计

$$\begin{aligned}&\|p_m\|_{L^\infty(0,T;H^1(\Omega))}+\|p_{kt}\|_{L^\infty(0,T;L^2(\Omega))}\\ \leqslant & C\Big(\|\mu\|_{\mathcal{M}([0,T];L^2(\Omega))}+\|g_y'(\cdot,y_m)\|_{L^1(0,T;L^2(\Omega))}+\big\|f_y'(\cdot,y_m(T))\big\|_{L^2(\Omega)}\Big).\end{aligned}$$

因为序列 $\{\lambda_m\}\subset\mathbb{R}$,

$$\{\mu_m\}\subset\mathcal{M}\big([0,T];L^2(\Omega)\big),\quad \{y_m\}\subset\mathcal{C}\big([0,T];L^2(\Omega)\big)$$

和 $\{u_m\}\subset L^2(\Sigma)$ 有界, 所以序列 $\{(p_m,p_{mt})\}$ 在 $L^\infty\big(0,T;H^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big)$ 中有界. 则存在

$$\begin{aligned}&(p_m,p_{mt})\overset{w^*}{\to}(p,p_t)\in L^\infty\big(0,T;H^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big),\\ &y_m\overset{w^*}{\to}\bar{y}\in L^\infty\big(0,T;L^2(\Omega)\big)\quad(m\to\infty),\end{aligned}$$

其中的收敛性都对应于所论空间的弱 * 拓扑. 另外知道 $u_m\to\bar{u}$ 在 $L^2(\Sigma)$ 的强拓扑下成立. 类似上述的标准推理, 易得 $\bar{y}$ 是 (7.27) 对应于 $\bar{u}$ 的解, 且 p 是 (7.30) 对应于 $\bar{y}$ 的 (唯一) 弱解.

现在证明极限 $(\lambda,\mu)=(\lambda,\bar{\mu}|_{]0,T]})$ 即为定理 7.8 断言了存在性的乘子. 首先根据 (H5) 中的凸性及集合 Θ 有非空内部的假设证明 $(\lambda,\mu)\neq 0$. 若不然, 则

$$\lim_{m\to\infty}|\mu_m|^2_{\mathcal{M}([0,T];L^2(\Omega))}=1.\tag{7.47}$$

根据假设 (H5) 有 $\widehat{y}_0\in\operatorname{int}\Theta$. 从而有闭球 $B_\rho(\widehat{y}_0)\subset\mathcal{C}\big([0,T];L^2(\Omega)\big)$ 完全包含于 Θ. 由 (7.47) 并任取 $m\in\mathbb{N}$, 可找到 $z_m\in\rho\mathbb{B}$ 满足

$$\langle z_m,\mu_m\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}([0,T];L^2(\Omega))}=\frac{\rho}{2}|\mu_m|_{\mathcal{M}([0,T];L^2(\Omega))}.$$

因为 $\widehat{y}_0+z_m\in\Theta$, 由 μ_m 的定义可以看到

$$\langle\widehat{y}_0+z_m-y_m,\mu_m\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}([0,T];L^2(\Omega))}\leqslant 0,\quad m\in\mathbb{N}.$$

在 $m\to\infty$ 时取极限得

$$\frac{\rho}{2}+\langle\widehat{y}_0-\bar{y},\bar{\mu}\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}([0,T];L^2(\Omega))}\leqslant 0.$$

已知 $\bar{y}(x,0)=\widehat{y}_0(x,0)$ 和 $\mu=\bar{\mu}|_{]0,T]}$, 因此

$$\langle\widehat{y}_0-\bar{y},\bar{\mu}\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}([0,T];L^2(\Omega))}=\langle\widehat{y}_0-\bar{y},\mu\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T];L^2(\Omega))},$$

这显然蕴涵

$$\langle\widehat{y}_0-\bar{y},\mu\rangle_{\mathcal{C}([0,T];L^2(\Omega)),\mathcal{M}_b(]0,T];L^2(\Omega))}\leqslant-\frac{\rho}{2}<0,$$

与假设 $(\lambda, \mu) = 0$ 矛盾, 从而就证明了 (7.28) 中的非平凡条件. 那里的第二个条件根据集合 Θ 的凸性也易由上述推理得到. 注意到 (7.28) 中的关系在总体上和约束优化问题的一般结果是一致的, 它们特别保证了约束具有 SNC 假设时的非平凡性; 比照命题 1.25 和下面的注 7.30.

剩下还需要证明最小值条件 (7.29). 为此注意到 $u_m \to \bar{u}$ ($L^2(\Sigma)$ 中强拓扑下), 在 (7.46) 中当 $m \to \infty$ 时取极限, 则有

$$\int_{\Sigma} \big(\lambda h(s,t,\bar{u}) + p\bar{u}\big)\,\mathrm{d}s\,\mathrm{d}t \leqslant \int_{\Sigma} \big(\lambda h(s,t,u_0) + pu_0\big)\,\mathrm{d}s\,\mathrm{d}t \tag{7.48}$$

对任意 $u_0 \in U_{ad}$ 成立. 最后注意到 (H6) 中容许控制集 U_{ad} 的结构, 并用标准的推理 (如 7.4 节中抛物方程的情形, 不需什么修改), 则可从 (7.48) 的积分条件得到逐点条件 (7.29). 证毕. △

注 7.21(双曲 Neumann 问题最优解的存在性)　简单起见, 本节没有涉及所论 Neumann 边界控制问题最优解的存在性问题. 但是, 由 7.2.2 小节给出的正则性结果可以导出该问题一个一般的存在性定理, 这基于下半连续函数在恰当拓扑下的紧集上极小化问题最优解存在性的 Weierstrass 定理; 比照后面的 7.3 节和 7.4 节, 那里有类似的推理和结果. 这里额外需要的是被积项 h 对于控制变量的凸性, 以及 (H6) 中控制集 $K(s,t)$ 的凸性. 这样的凸性对逐点必要最优条件并不必要, 但存在性定理却需要它们来保证费用函数的下半连续性和可行控制集相对于控制弱收敛的闭性, 进而由正则性可导出轨道的强收敛. 注意对下节要考虑的 Dirichlet 边界控制问题, 则需要假设苛刻得多的条件, 控制变量和状态变量的凸性对最优解的必要最优条件和存在性都是需要的. 这是因为, Dirichlet 边界控制问题和 Neumann 类型的对应问题相比正则性要少得多, 这包括双曲和抛物的情形, 所以它需要变分分析的不同方法; 见 7.3 节和 7.4 节.

7.3　线性约束双曲方程的 Dirichlet 边界控制

本节研究 7.2 节讨论过的具有逐点状态约束的双曲方程 Neumann 边界控制问题的 Dirichlet 版本. 如前所述, 对双曲方程, Neumann 和 Dirichlet 边界条件差别是很大的, 因此本节建立的方法和结果与 7.2 节也相当不同. 粗略地讲, 对 Dirichlet 问题初始数据加的条件要强, 但相比之下所得的结果却要弱. 这是由于 Dirichlet 情形正则性的缺失, 故这里不得不对其的变分分析发展不同的方法. 特别地, 必要最优条件是通过把动态优化的 Dirichlet 控制问题化归为具有特殊类型的几何和算子约束的无穷维线性规划问题来实现的.

7.3.1 Dirichlet 控制问题的表述和主要结果

下面的讨论中保持 7.2 节的标准记号. 所考虑的 Dirichlet 问题表述如下. 给定有界开定义域 $\Omega \subset \mathbb{R}^n$, 其边界 Γ 为 $\mathcal{C}^2$ 类型, 考虑极小化积分泛函

$$J(y,u) := \int_\Omega f\big(x, y(T)\big)\,\mathrm{d}x + \int_Q g(x,t,y)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma h(s,t,u)\,\mathrm{d}s\,\mathrm{d}t,$$

其中 $T>0$ 是固定时间, 而容许解对 $\{y(\cdot),u(\cdot)\}$ 满足高维线性波动方程, 控制函数作用于"Dirichlet 边界条件":

$$\begin{cases} y_{tt} - \Delta y = \vartheta, & \text{在 } Q := \Omega\times(0,T) \text{ 中},\\ y = u, & \text{在 } \Sigma := \Gamma\times(0,T) \text{ 中},\\ y(0)=y_0, \quad y_t(0) = y_1, & \text{在 } \Omega \text{ 中}, \end{cases} \tag{7.49}$$

并满足逐点控制和状态约束

$$u(\cdot)\in U_{ad}\subset L^2(\Sigma), \quad y(\cdot)\in\Theta\subset\mathcal{C}\big([0,T];L^2(\Omega)\big),$$

其中 $\vartheta\in L^1\big(0,T;H^{-1}(\Omega)\big)$, $y_0\in L^2(\Omega)$ 和 $y_1\in H^{-1}(\Omega)$ 为给定的函数. 记此问题为 (DP) 并简写为

$$\inf\Big\{J(y,u)\mid\{y(\cdot),u(\cdot)\}\text{ 满足(7.49)}, u(\cdot)\in U_{ad}, y(\cdot)\in\Theta\Big\}.$$

本节的基本目标是导出所论的 Dirichlet 状态约束问题 (DP) 的必要最优条件, 这与 7.2 节讨论的 Neumann 问题 (NP) 一样. 但是, 与 (NP) 相比, 这里必须对 (DP) 的初始数据加更严苛的条件, 而得到的结果却更弱; 见下述. 注意 (DP) 中的双曲动态系统是由线性波动方程描述的, ϑ 不依赖于 y, 这与 (NP) 中的半线性不同. 另一方面, 这里对 Dirichlet 问题的初始状态只要求更弱的 $(y_0,y_1)\in L^2(\Omega)\times H^{-1}(\Omega)$, 而不是 Neumann 情形的 $(y_0,y_1)\in H^1(\Omega)\times L^2(\Omega)$. 事实上, 对 (DP) 得到的结果可以推广到更一般的线性双曲方程的情形, 即由一个强椭圆算子代替 Laplace 算子 Δ.

现在描述对 (DP) 中初始数据的常驻假设以便推导后面的必要最优条件; 只有前四个假设在后面的存在性定理中是需要的 (此时 (H4) 中不需要假设 $\operatorname{int}\Theta\neq\varnothing$).

(H1) 对任意 $y\in\mathbb{R}$, 函数 $f(\cdot,y)\geqslant 0$ 在 Ω 上可测且 $f(\cdot,0)\in L^1(\Omega)$. 对几乎处处的 $x\in\Omega$, 函数 $f(x,\cdot)$ 在整个实直线 $\mathbb{R}$ 上连续凸.

(H2) 对任意 $y\in\mathbb{R}$, 函数 $g(\cdot,\cdot,y)\geqslant 0$ 在 Q 上可测且 $g(\cdot,\cdot,0)\in L^1(Q)$. 对几乎处处的 $(x,t)\in Q$, 函数 $g(x,t,\cdot)$ 在 $\mathbb{R}$ 上连续凸.

(H3) 对任意 $u\in\mathbb{R}$, 函数 $h(\cdot,u)$ 在 Σ 上可测且 $h(\cdot,0)\in L^1(\Sigma)$. 对几乎处处的 $(s,t)\in\Sigma$, 函数 $h(s,t,\cdot)$ 在 $\mathbb{R}$ 上连续凸. 进一步, h 满足如下增长条件

$$|u|^2\leqslant h(s,t,u), \quad \forall(s,t)\in\Sigma, \quad u\in\mathbb{R}.$$

(**H4**) 状态约束集 $\Theta \subset \mathcal{C}\big([0,T];L^2(\Omega)\big)$ 是闭凸的且 $\operatorname{int}\Theta \neq \varnothing$. 控制集 $U_{ad} \subset L^2(\Sigma)$ 也是闭凸的. 进一步, 对初始函数 $(x,t) \mapsto y_0(x)$ 有 $y_0(\cdot) \in \operatorname{int}\Theta$, 并且存在 $u \in U_{ad}$ 满足 $y_u \in \Theta$ 和 $J(y,u) < \infty$, 其中 $y(\cdot)$ 为 Dirichlet 系统 (7.49) 的解.

(**H5**) 对几乎处处的 $x \in \Omega$, 函数 $f(x,\cdot)$ 是 $\mathcal{C}^1$ 类型的且

$$|f_y'(x,y)| \leqslant C(1+|y|)$$

对某常数 $C>0$ 成立.

(**H6**) 对几乎处处的 $(x,t) \in Q$, 函数 $g(x,t,\cdot)$ 是 $\mathcal{C}^1$ 类型的且

$$|g_y'(x,t,y)| \leqslant C(1+|y|)$$

对某常数 $C>0$ 成立.

(**H7**) 对几乎处处的 $(s,t) \in \Sigma$, 函数 $h(s,t,\cdot)$ 是 $\mathcal{C}^1$ 类型的且

$$|h_u'(s,t,u)| \leqslant C(1+|u|)$$

对某常数 $C>0$ 成立.

与 (NP) 相比, 对 (DP) 的假设的主要区别在于被积项 f,g,h 对于状态和控制变量的完整凸性和控制集 U_{ad} 的凸性, 这在 Neumann 问题中是没有的. 如前述, 这是因为 (DP) 没有足够的正则性; 更多细节和讨论见 7.3.2 小节.

事实上, 额外的凸性假设在某中意义上弥补了正则性的缺失. 基于完整的凸性和所具有的正则性, 可以把该 Dirichlet 控制问题化归为 Banach 空间上具有几何和算子约束的特殊数学规划问题, 然后由恰当版本的数学规划 (抽象)Lagrange 乘子法则 (见 5.1.2 小节) 导出 (DP) 的必要最优条件. 由此得出的最优条件具有 Pontryagin 最大值原理的积分形式, 这与 7.2 节中 Neumann 问题情形的逐点形式是不同的. 进一步, 所给假设能允许建立问题 (DP) 最优控制的一个一般存在性定理.

现在表述本节的主要结果. 这些结果需要的状态和伴随方程的 (弱) 解等恰当术语会分别在 7.3.3 小节和 7.3.4 小节中严格澄清.

定理 7.22(Dirichlet 最优控制的存在性) 设 (H1)~(H4) 都成立 (在 (H4) 中不需要 $\operatorname{int}\Theta \neq \varnothing$), 则 Dirichlet 最优控制问题 (DP) 有最优解.

证明在 7.3.2 中给出.

定理 7.23(双曲 Dirichlet 问题的必要最优条件) 设 (H1)~(H7) 都成立. 则对 (DP) 的任一最优解 $\{\bar{y}(\cdot),\bar{u}(\cdot)\}$ 下述条件成立: 存在 $\lambda \geqslant 0$ 和 $\mu \in \mathcal{M}_b\big(]0,T];L^2(\Omega)\big)$ 使得

$$(\lambda,\mu) \neq 0, \quad \langle \mu, y-\bar{y}\rangle \leqslant 0, \quad \forall y \in \Theta, \tag{7.50}$$

$$\int_\Sigma \Big(\frac{\partial p}{\partial \nu} + \lambda h_u'(s,t,\bar{u})\Big)(u-\bar{u})\,\mathrm{d}s\,\mathrm{d}t \geqslant 0, \quad \forall u \in U_{ad}, \tag{7.51}$$

其中 p 是对应的如下伴随系统的解:

$$\begin{cases} p_{tt} - \Delta p = \lambda g_y'(x,t,\bar{y}) + \mu|_Q, & \text{在 } Q \text{ 中}, \\ p = 0, & \text{在 } \Sigma \text{ 中}, \\ p(T) = y_0, \quad p_t(T) = -\lambda f_y'\big(x, \bar{y}(T)\big) - \mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中}. \end{cases} \tag{7.52}$$

进一步, 如果存在 $\{y(\cdot), u(\cdot)\} \in Y \times (U_{ad} - \bar{u})$ 满足

$$\begin{cases} y_{tt} - \Delta y = 0, & \text{在 } Q \text{ 中}, \\ y = u, & \text{在 } \Sigma \text{ 中}, \\ y(0) = 0, \quad y_t(0) = 0, & \text{在 } \Omega \text{ 中}, \quad \bar{y} + y \in \operatorname{int}\Theta, \end{cases} \tag{7.53}$$

其中状态空间 Y 定义于 (7.54), 则可在上述最优条件中取 $\lambda = 1$.

注意积分条件 (7.51) 是含于一个"最小值原理"(而非最大值原理) 中的, 在本节的框架下这是更方便的. 定理 7.23 的证明在 7.3.4 小节给出, 而伴随系统的预备分析在将 7.3.3 小节给出.

7.3.2 Dirichlet 最优控制的存在性

对下面的讨论首先需要重述非齐次 Dirichlet 状态系统 (7.49) 的一个恰当的解的概念. 下述的"弱解"概念满足这里的要求.

定义 7.24(Dirichlet 状态双曲系统的弱解) 满足 $(y, y_t) \in \mathcal{C}\big([0,T]; L^2(\Omega)\big) \times \mathcal{C}\big([0,T]; H^{-1}(\Omega)\big)$ 的函数 $y(\cdot)$ 称为 (7.49) 的一个"弱解", 如果

$$\begin{aligned} \int_Q fz \,\mathrm{d}x\,\mathrm{d}t = & \int_Q y\varphi \,\mathrm{d}x\,\mathrm{d}t + \big\langle y_t(T), z^0 \big\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} \\ & - \big\langle y_t(0), z(0) \big\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} \\ & - \int_\Omega y(T) z^1 \,\mathrm{d}x + \int_\Omega y(0) z_t(0) \,\mathrm{d}x + \int_\Sigma \frac{\partial z}{\partial \nu_u} \,\mathrm{d}s\,\mathrm{d}t \end{aligned}$$

对任意 $(\varphi, z^0, z^1) \in L^1\big(0,T; L^2(\Omega)\big) \times H_0^1(\Omega) \times L^2(\Omega)$ 成立, 其中 z 是下述齐次 Dirichlet 问题的解:

$$\begin{cases} z_{tt} - \Delta z = \varphi, & \text{在 } Q \text{ 中}, \\ z = 0, & \text{在 } \Sigma \text{ 中}, \\ z(T) = z^0, \quad z_t(T) = z^1, & \text{在 } \Omega \text{ 中}. \end{cases}$$

这里定义的弱解概念的重要性源于如下基本正则性结果, 它由 Lasiecka, Lions, 和 Triggiani 在文献 [740] 中建立, 在恰当的 Banach 空间中这保证了 (7.49) 弱解的存在性、唯一性, 及其对初始和边界条件的连续依赖性.

定理 7.25(Dirichlet 双曲问题的基本正则性) 对任意 $(\vartheta, u, y_0, y_1) \in L^1\big(0,T; H^{-1}(\Omega)\big) \times L^2(\Sigma) \times L^2(\Omega) \times H^{-1}(\Omega)$, Dirichlet 系统 (7.49) 存在唯一满足 $(y, y_t) \in \mathcal{C}\big([0,T]; L^2(\Omega)\big) \times \mathcal{C}\big([0,T]; H^{-1}(\Omega)\big)$ 的弱解 $y(\cdot)$. 进一步, 映射 $(\vartheta, u, y_0, y_1) \mapsto (y, y_t)$ 从 $L^1\big(0,T; H^{-1}(\Omega)\big) \times L^2(\Sigma) \times L^2(\Omega) \times H^{-1}(\Omega)$ 到 $\mathcal{C}\big([0,T]; L^2(\Omega)\big) \times \mathcal{C}\big([0,T]; H^{-1}(\Omega)\big)$ 是线性连续的.

定理 7.25 在进一步的讨论中至关重要. 受该定理启发, 下面引入"容许状态函数"空间, 即系统 (7.49) 在 $(\vartheta, u, y_0, y_1) \in L^1\big(0,T; H^{-1}(\Omega)\big) \times L^2(\Sigma) \times L^2(\Omega) \times H^{-1}(\Omega)$ 时的解空间, 其定义如下:

$$\begin{aligned} Y := \Big\{ & y \in \mathcal{C}\big([0,T]; L^2(\Omega)\big) \mid y_t \in \mathcal{C}\big([0,T]; H^{-1}(\Omega)\big), \\ & y_{tt} - \Delta y \in L^1\big(0,T; H^{-1}(\Omega)\big), \quad y|_\Sigma \in L^2(\Sigma) \Big\}. \end{aligned} \tag{7.54}$$

易见 Y 为 Banach 空间, 其中范数 $\|\cdot\|$ 定义为

$$\|y\|_{\mathcal{C}([0,T];L^2(\Omega))} + \|y_t\|_{\mathcal{C}([0,T];H^{-1}(\Omega))} + \|y_{tt} - \Delta y\|_{L^1(0,T;H^{-1}(\Omega))} + \|y|_\Sigma\|_{L^2(\Sigma)}.$$

基于定理 7.25 和在所给假设以及恰当弱拓扑下关于积分函数下半连续性的标准结果, 下面证明 (DP) 的最优解的存在性, 方法是将其化归为相关空间上经典的 Weierstrass 定理.

定理 7.22 的证明 由定理 7.25 中的存在性和唯一性表述, 存在问题 (DP) 中的极小化序列 $\{(y_k, u_k)\} \subset \mathcal{C}\big([0,T]; L^2(\Omega)\big) \times U_{ad}$, 其中 y_k 是 (7.49) 相对于 u_k 的 (唯一) 解. 根据 (H3) 中的增长性条件, 序列 $\{u_k\}$ 在 $L^2(\Sigma)$ 中是有界的. 因此, 不失一般性, 可设 $\{u_k\}$ 在 $L^2(\Sigma)$ 中沿弱拓扑收敛到 u. 因为 (H4) 中假设 U_{ad} 是闭凸的, 则 $u(\cdot) \in U_{ad}$. 由定理 7.25 的连续性表述, 序列 $\{(y_k, y_{kt})\}$ 在 $L^\infty\big(0,T; L^2(\Omega)\big) \times L^\infty\big(0,T; H^{-1}(\Omega)\big)$ 中是有界的, 其中 y_{kt} 表示 y_k 的分布导数. 根据上面的连续性, $\{(y_k, y_{kt})\}$ 在 $L^\infty\big(0,T; L^2(\Omega)\big) \times L^\infty\big(0,T; H^{-1}(\Omega)\big)$ 中沿弱 * 拓扑收敛到 (y, y_t), 其中 y 是 (7.49) 相对于 u 的解. 由 (H4) 中给出的 Θ 的闭凸性, 则得 $y(\cdot) \in \Theta$.

剩下来需要证明费用函数的下半连续性, 即证极限关系

$$J(y, u) \leqslant \liminf_{k\to\infty} J(y_k, u_k).$$

由假设 (H1)~(H3) 中关键的凸性假设, 此式可直接由积分函数对于所论弱拓扑下半连续性的经典结果得到. 因此 (y, u) 是 Dirichlet 最优控制问题 (DP) 的一个最优解. △

7.3.3 Dirichlet 问题中的伴随系统

本小节的主要目标是证明定理 7.23 表述的必要最优条件. 为此先澄清该定理中伴随系统解的恰当概念, 然后建立伴随轨道的一些性质, 从而由 Banach 空间中

的辅助优化问题的抽象必要最优条件导出所需的双曲控制问题的必要最优条件. 给定 $\mu \in \mathcal{M}_b(]0,T];L^2(\Omega))$, 考虑对应于定理 7.23 中伴随系统 (7.52) (在 $(\lambda, y_0) = 0$ 的情形) 的系统

$$\begin{cases} p_{tt} - \Delta p = \mu|_Q, & \text{在 } Q \text{ 中}, \\ p = 0, & \text{在 } \Sigma \text{ 中}, \\ p(T) = 0, \quad p_t(T) = -\mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中}, \end{cases} \tag{7.55}$$

其中 $\mu|_Q$ (对应地, $\mu|_{\Omega\times\{T\}}$) 为 μ 在 Q (对应地, $\Omega\times\{T\}$) 上的限制. 可以看到, 这些限制满足 $\mu|_Q \in \mathcal{M}_b(]0,T[;L^2(\Omega))$ 和 $\mu|_{\Omega\times\{T\}} \in L^2(\Omega)$.

为恰当定义伴随系统 (7.55) 的解, 临时假设 $(p,p_t) \in L^2(0,T;H_0^1(\Omega)) \times L^2(0,T;L^2(\Omega))$ 和 $p_{tt} - \Delta p \in \mathcal{M}_b(]0,T[;L^2(\Omega))$, 其中导数在 Q 中分布的意义下计算. 基于引理 7.14 中的散度公式, 沿着 7.2.2 小节中 Neumann 问题的证明路线, 定义向量场 $(-\nabla p, p_t)$ 在 ∂Q 上的法向迹, 它是 $H^{-1/2}(\partial Q)$ 中的元, 记为 $\gamma_{\nu_Q}(-\nabla p, p_t)$. 则有估计

$$\begin{aligned} &\|\gamma_{\nu_Q}(-\nabla p, p_t)\|_{H^{-1/2}(\partial Q)} \\ \leqslant\, & C\big(\|p\|_{L^2(0,T;H_0^1(\Omega))} + \|p_t\|_{L^2(\Omega)} + \|p_{tt} - \Delta p\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))}\big), \end{aligned}$$

其中 C 独立于 p. 这就可以定义 $p_t(0)$ 为该法向迹在 $\Omega\times\{0\}$ 上的限制, 即

$$\gamma_{\nu_Q}(-\nabla p, p_t)|_{\Omega\times\{0\}} = p_t(0) \in H^{-1/2}(\Omega).$$

由此有伴随系统 (7.55) 的如下弱解定义.

定义 7.26(Dirichlet 伴随系统的弱解) 满足 $(p,p_t) \in L^\infty(0,T;H_0^1(\Omega)) \times L^\infty(0,T;L^2(\Omega))$ 和 $p_{tt} - \Delta p \in \mathcal{M}_b(]0,T[;L^2(\Omega))$ 的函数 p 称为 Dirichlet 伴随系统 (7.55) 的一个"弱解", 如果等式

$$\begin{aligned} &-\int_\Omega p(0)y_1 \,\mathrm{d}x + \langle p_t(0), y_0\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} \\ &+\langle y(\vartheta, y_0, y_1), \mu\rangle_{C([0,T];L^2(\Omega))\times\mathcal{M}_b(]0,T];L^2(\Omega))} - \int_Q p\vartheta \,\mathrm{d}x\,\mathrm{d}t = 0 \end{aligned}$$

对任意 $(\vartheta, y_0, y_1) \in L^2(\Omega) \times H_0^1(\Omega) \times L^2(\Omega)$ 成立, 其中 $y(\vartheta, y_0, y_1)$ 表示 (7.49) 中齐次 Dirichlet 问题 (如下) 的唯一解:

$$\begin{cases} y_{tt} - \Delta y = \vartheta, & \text{在 } Q \text{ 中}, \\ y = 0, & \text{在 } \Sigma \text{ 中}, \\ y(0) = y_0, \quad y_t(0) = y_1, & \text{在 } \Omega \text{ 中}. \end{cases}$$

可以看出, 因为 $(p,p_t)\in L^\infty\big(0,T;H_0^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big)$, 所以 $p\in\mathcal{C}\big([0,T];L^2(\Omega)\big)$, 从而 $\int_\Omega p(0)y_1\,\mathrm{d}x$ 是有意义的. 进一步, $p_{tt}-\Delta p\in\mathcal{M}_b\big(]0,T[;L^2(\Omega)\big)$, 所以根据上面定义之前的讨论, $p_t(0)=\gamma_{\nu_Q}(-\nabla p,p_t)|_{\Omega\times\{0\}}$ 在 $H^{-1/2}(\Omega)$ 中也是有意义的.

下面这个重要的结果证明了定义 7.26 意义下伴随系统 (7.55) 的弱解的存在性和唯一性, 另外还对主要定理的证明提供了很有意义的额外的正则性性质.

定理 7.27(Dirichlet 问题伴随轨线的性质) 伴随系统 (7.55) 存在一个唯一弱解 $p(\cdot)$ 满足

$$(p,p_t)\in L^\infty\big(0,T;H_0^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big).$$

此外, 有

$$\begin{aligned}
&p_t\in BV\big([0,T];H^{-1}(\Omega)\big),\\
&\frac{\partial p}{\partial\nu}=\gamma_{\nu_Q}(\nabla p,-p_t)|_\Sigma\in L^2(\Sigma),\\
&p\in\mathcal{C}_w\big([0,T];H_0^1(\Omega)\big),\\
&p_t(\tau)\in L^2(\Omega),\quad \forall\tau\in\big\{t\in[0,T]\mid\mu(\Omega\times\{t\})=0\big\}.
\end{aligned}$$

特别地, $p_t(0)\in L^2(\Omega)$.

证明 首先注意到若对 $(p,p_t)\in L^\infty\big(0,T;H_0^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big)$ 在 $\mu=0$ 时定理中的等式成立, 则显然 $p=0$. 所以 (7.55) 至多有一个定义 7.26 意义下的解. 因此只需要再在定理的额外正则性假设下证明弱解的存在性就可以了.

令 $\{\mu_k\}$ 为 $L^1\big(0,T;L^2(\Omega)\big)$ 中满足

$$\begin{aligned}
&\|\mu_k\|_{L^1(0,T;L^2(\Omega))}=\|\mu|_{]0,T[}\|_{\mathcal{M}_b(]0,T[;L^2(\Omega))},\\
&\lim_{k\to\infty}\int_Q y\mu_k\,\mathrm{d}x\,\mathrm{d}t=\langle y,\mu|_{]0,T[}\rangle_{\mathcal{C}([0,T];L^2(\Omega))\times\mathcal{M}_b(]0,T[;L^2(\Omega))}\\
&\qquad\qquad\big(\text{若}\,y\in\mathcal{C}\big([0,T];L^2(\Omega)\big)\big)
\end{aligned}$$

的序列. 记 p_k 为下述系统的 (唯一) 解:

$$\begin{cases}
p_{tt}-\Delta p=\mu_k, & \text{在 } Q \text{ 中},\\
p=0, & \text{在 } \Sigma \text{ 中},\\
p(T)=0,\quad p_t(T)=-\mu|_{\Omega\times\{T\}}, & \text{在 } \Omega \text{ 中}.
\end{cases}\tag{7.56}$$

应用前面提到的 Lasiecka, Lions 和 Triggiani 的文章 [740] 中的定理 2.1, 则有估计

$$\begin{aligned}
&\|p_k\|_{L^\infty(0,T;H_0^1(\Omega))}+\|p_{kt}\|_{L^\infty(0,T;L^2(\Omega))}+\Big\|\frac{\partial p_k}{\partial\nu}\Big\|_{L^2(\Sigma)}\\
&\quad+\|p_k(0)\|_{H^1(\Omega)}+\|p_{kt}(0)\|_{L^2(\Omega)}\leqslant C\|\mu\|_{\mathcal{M}_b(]0,T];L^2(\Omega))},
\end{aligned}$$

其中常数 C 与 k 无关. 由 (7.56), p_{kt} 关于 t 的分布导数可表示为形式

$$
\begin{aligned}
p_{ktt} = \pi_k + \mu_k \in L^\infty\big(0,T;H^{-1}(\Omega)\big) + \mathcal{M}_b\big(]0,T[;L^2(\Omega)\big) \\
\subset \mathcal{M}_b\big(]0,T[;H^{-1}(\Omega)\big),
\end{aligned}
$$

其中 π_k 定义为

$$
\big\langle \pi_k, y\big\rangle_{L^\infty(0,T;H^{-1}(\Omega))\times L^1(0,T;H_0^1(\Omega))} := -\int_Q \nabla p_k \nabla y \,\mathrm{d}x\,\mathrm{d}t.
$$

所以 $\{p_{ktt}\}$ 在 $\mathcal{M}_b\big(]0,T[;H^{-1}(\Omega)\big)$ 中有界, 从而对应的 $\{p_{kt}\}$ 在 $BV\big([0,T];H^{-1}(\Omega)\big)$ 中有界. 于是存在 $p \in L^\infty\big(0,T;H_0^1(\Omega)\big)$ 满足 $p_t \in BV\big([0,T];H^{-1}(\Omega)\big)$, 以及 $\{p_k\}$ 的子序列满足 $p_k \to p$ (沿 $L^\infty\big(0,T;H_0^1(\Omega)\big)$ 的弱 * 拓扑) 和 $p_{kt} \to p_t$ (沿 $L^\infty\big(0,T;L^2(\Omega)\big)$ 的弱 * 拓扑), 其中 $k \to \infty$. 因为序列 $\{\gamma_{\nu_Q}(-\nabla p_k, p_{kt})\}$ 在 $L^2(\partial Q)$ 中有界, 可以假设弱收敛

$$
\gamma_{\nu_Q}(-\nabla p_k, p_{kt}) \to \gamma_{\nu_Q}(-\nabla p, p_t) \quad (\text{沿} L^2(\partial Q) \text{ 的弱拓扑}).
$$

另一方面, $\gamma_{\nu_Q}\big(-\nabla p_k, p_{kt}\big)|_{\Omega\times\{T\}} = \mu|_{\Omega\times\{T\}}$ 且序列

$$
\gamma_{\nu_Q}\big(\nabla p_k - p_{kt}\big)|_\Sigma = \frac{\partial p_k}{\partial \nu}
$$

在 $L^2(\Sigma)$ 中有界. 因此

$$
\begin{aligned}
&\gamma_{\nu_Q}\big(\nabla p_k, -p_{kt}\big)|_\Sigma \to \gamma_{\nu_Q}\big(\nabla p, -p_t\big)|_\Sigma = \frac{\partial p}{\partial \nu}, \\
&\gamma_{\nu_Q}\big(-\nabla p_n, p_{nt}\big)|_{\Omega\times\{0\}} = p_{nt}(0) \to \gamma_{\nu_Q}\big(-\nabla p, p_t\big)|_{\Omega\times\{0\}} = p_t(0)
\end{aligned}
$$

分别在 $L^2(\Sigma)$ 和 $L^2(\Omega)$ 的弱拓扑下成立. 在 $k \to \infty$ 时在等式

$$
\begin{aligned}
&-\big\langle p_k(0), y_1\big\rangle_{L^2(\Omega)} + \big\langle p_{kt}(0), y_0\big\rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} \\
&+\big\langle y(\vartheta, y_0, y_1), \mu_k\big\rangle_{\mathcal{C}([0,T];L^2(\Omega))\times\mathcal{M}_b(]0,T];L^2(\Omega))} - \big\langle p_k, \vartheta\big\rangle_{L^2(\Omega)} = 0
\end{aligned}
$$

中取极限, 则得 $p(\cdot)$ 为伴随系统 (7.55) 满足定理中给出的除最后一个外所有其他关系的弱解.

为证最后一个关系, 假设 $\mu(\Omega\times\{t\}) = 0$ 对某 $t \in [0,T]$ 成立. 如前考虑 $(-\nabla p, p_t)$ 在 $\partial(\Omega\times]0,t[)$ 上的法向迹, 就得等式

$$
\gamma_{\nu_{\Omega\times]0,t[}}\big(-\nabla p, p_t\big)|_{\Omega\times\{t\}} = p_t(0) \in L^2(\Omega).
$$

这就完成了定理的证明. △

本节的最后给出定理 7.27 在极限情形的一个有用推论, 它给出了伴随系统 (7.55) 的解和容许状态函数空间 Y (见 (7.54)) 中的原始轨线的 Green 类型的关系.

定理 7.28(Dirichlet 双曲问题的 Green 公式)　对给定的测度 $\mu \in \mathcal{M}_b(]0,T]; L^2(\Omega))$ 考虑伴随系统 (7.55) 的唯一解 $p(\cdot)$. 则对任意容许状态函数 $y \in Y$, 伴随轨线 $p(\cdot)$ 满足下述 Green 公式:

$$
\begin{aligned}
&\langle y, \mu \rangle_{\mathcal{C}([0,T];L^2(\Omega))\times\mathcal{M}_b(]0,T];L^2(\Omega))} - \langle p, y_{tt} - \Delta y \rangle_{L^\infty(0,T;H_0^1(\Omega))\times L^1(0,T;H^{-1}(\Omega))} \\
= &-\int_\Omega y(0)p_t(0)\,\mathrm{d}x + \langle y_t(0), p(0) \rangle_{H^{-1}(\Omega)\times H_0^1(\Omega)} - \int_\Sigma y\frac{\partial p}{\partial \nu}\,\mathrm{d}s\,\mathrm{d}t.
\end{aligned}
$$

证明　如定理 7.27 中建立的, 上述 Green 公式对逼近伴随系统 (7.56) 的解 p_k 成立. 对 $k \to \infty$ 取极限即得所需结果. △

7.3.4　最优条件的证明

本小节证明定理 7.23. 所用策略如下:

(a) 将 (DP) 化为 Banach 空间中的一个一般的具有几何和算子约束的优化/数学规划问题, 其广义 Lagrange 乘子型的必要最优条件是已知的.

(b) 用原 Dirichlet 控制问题 (DP) 的初始数据来表述这些必要最优条件及相关的假设.

除了用到第 5 章中的一般优化理论, 该证明本质上基于 7.3.3 小节得到的所论双曲系统的具体结果, 而这些结果则用到了定理 7.25 中的正则性. 把 (DP) 化为的这个 Banach 空间上的一般优化问题, 方便起见称为"抽象控制问题", 表述如下:

$$
\begin{aligned}
&\min \quad \varphi(z,w) \\
&\text{s.t.} \quad z \in Z, \quad w \in W_{ad}, \quad f_1(z,w) = 0, \quad f_2(w) \in \Xi,
\end{aligned}
\tag{7.57}
$$

其中 $\varphi\colon Z \times W \to \mathbb{R}$, $f_1\colon Z \times W \to Z_1$, $f_2\colon Z \to Z_2$, 集合 $W_{ad} \subset W$ 和 $\Xi \subset Z_2$ 为闭凸的, 空间 Z, Z_1, Z_2, W 为 Banach 空间, 而且 W 是可分的.

注意到, "抽象控制问题" (7.57) 是 5.1.2 小节研究过的具有几何和算子约束的数学规划类型. 下面假设费用函数的"Fréchet可微性"和约束算子的"严格可微性", 且严格导数是满射. 再计入问题 (7.57) 的特殊结构和描述几何约束的集合的凸性, 就能得出 (7.57) 的必要最优条件, 这可作为定理 5.11(ii) 细致的补充, 因为在所给的光滑性和凸性假设下, 这里的定义域空间可以是任意 Banach 空间. 下述定理在具有光滑算子约束凸优化的观点下的一个直接证明见 Alibert 与 Raymond [9].

定理 7.29(抽象控制问题的必要条件)　设 $(\bar{z}, \bar{w})$ 为问题 (7.57) 的一个最优解, 并设 φ 在 $(\bar{z}, \bar{w})$ 点 Fréchet可微, f_1 和 f_2 分别在 $\bar{z}$ 和 $(\bar{z}, \bar{w})$ 严格可微, 偏导数

$f'_{1z}(\bar z,\bar w)\colon Z\to Z_1$ 为满射, 且 $\operatorname{int}\Xi\neq\varnothing$. 则存在伴随元 $(p,\mu,\lambda)\in Z_1^*\times Z_2^*\times\mathbb{R}^+$ 满足 $(\lambda,\mu)\neq 0$ 使得下述条件成立:

$$
\begin{aligned}
&\lambda\varphi'_z(\bar z,\bar w)z+\big\langle p,f'_{1z}(\bar z,\bar w)z\big\rangle+\big\langle \mu,f'_2(\bar z)z\big\rangle=0,\quad \forall z\in Z,\\
&\big\langle \mu,z-f_2(\bar z)\big\rangle\leqslant 0,\quad \forall z\in\Xi,\\
&\lambda\varphi'_w(\bar z,\bar w)(w-\bar w)+\big\langle p,f'_{1w}(\bar z,\bar w)(w-\bar w)\big\rangle\geqslant 0,\quad \forall w\in W_{ad}.
\end{aligned}
$$

若进一步假设

$$
f'_{1z}(\bar z,\bar w)z_0+f'_{1w}(\bar y,\bar w)w_0=0,\quad f_2(\bar z)+f'_2(\bar z)z_0\in\operatorname{int}\Xi
$$

对某 $w_0\in(W_{ad}-\bar w)$ 和 $z_0\in Z$ 成立, 则上述条件可取成标准/规范形式, 即可取 $\lambda=1$.

本节最后证明所表述的原 Dirichlet 控制问题 (DP) 的必要最优条件.

定理 7.23 的证明 设 $(\bar y,\bar u)\in Y\times U_{ad}$ 为 (DP) 的一个参考最优解. 现在把 (DP) 化为定理 7.29 中考虑的线性规划问题 (7.57). 置

$$
Z:=Y,\quad (z,w):=(y,u),\quad W:=L^2(\Sigma),\quad W_{ad}:=U_{ad},\quad \Xi:=\Theta,
$$

$$
Z_1:=L^1\big(0,T;H^{-1}(\Omega)\big)\times L^2(\Sigma)\times L^2(\Omega)\times H^{-1}(\Omega),\quad Z_2:=\mathcal{C}\big([0,T];L^2(\Omega)\big),
$$

$$
\varphi(y,u):=J(y,u),\quad f_1(y,u):=\big(y_{tt}-\Delta y-\vartheta,y|_\Sigma-u,y(0)-y_0,y_t(0)-y_1\big),
$$

且 $f_2(y):=y$. 由假设 (H5)~(H7) 知, 费用函数 φ 在 $(\bar y,\bar u)$ 点 Fréchet可微, 映射 f_1 在 $(\bar y,\bar u)$ 点严格可微, 且

$$
\begin{aligned}
\varphi'(\bar y,\bar u)(y,u)=&\int_\Omega f'_y\big(x,\bar y(T)\big)y(T)\,\mathrm{d}x+\int_Q g'_y(x,t,\bar y)y\,\mathrm{d}x\,\mathrm{d}t\\
&+\int_\Sigma h'_u(s,t,\bar u)u\,\mathrm{d}s\,\mathrm{d}t\\
f'_1(\bar y,\bar u)(y,u)=&f'_{1y}(\bar y,\bar u)y+f'_{1u}(\bar y,\bar u)u,\\
f'_{1y}(\bar y,\bar u)y=&\big(y_{tt}-\Delta y,y|_\Sigma,y(0),y_t(0)\big),\\
f'_{1u}(\bar y,\bar u)=&(0,-u,0,0),\quad \forall (y,u)\in Y\times L^2(\Sigma).
\end{aligned}
$$

进一步, 由定理 7.25 得线性连续算子 $f'_{1y}(\bar y,\bar u)$ 是从 Y 到 $L^1\big(0,T;H^{-1}(\Omega)\big)\times L^2(\Sigma)\times L^2(\Omega)\times H^{-1}(\Omega)$ 的满射. 因此定理 7.29 的所有条件都满足了.

由此定理找到 $\lambda\in\mathbb{R}^+$, $(\bar p,\widetilde p,\widehat p,\breve p)\in L^\infty\big(0,T;H_0^1(\Omega)\big)\times L^2(\Sigma)\times L^2(\Omega)\times H_0^1(\Omega)$ 和 $\mu\in\mathcal{M}\big([0,T];L^2(\Omega)\big)$ 满足 $(\lambda,\mu)\neq 0$ 使得

$$\begin{aligned}&\int_{\Omega}\lambda f_y'\big(x,\bar{y}(T)\big)y(T)\,\mathrm{d}x+\int_{Q}\lambda g_y'(x,t,\bar{y})y\,\mathrm{d}x\,\mathrm{d}t+\big\langle\bar{p},y_{tt}-\Delta y\big\rangle\\&+\int_{\Sigma}\widetilde{p}\,y\,\mathrm{d}s\,\mathrm{d}t+\big\langle\widehat{p},y(0)\big\rangle+\big\langle\breve{p},y_t(0)\big\rangle\\&+\big\langle\mu,y\big\rangle_{\mathcal{M}([0,T];L^2(\Omega))\times\mathcal{C}([0,T];L^2(\Omega))}=0\end{aligned}\tag{7.58}$$

对 (7.54) 中容许状态函数空间 Y 中的任意 y 成立, 且

$$\big\langle\mu,z-\bar{y}\big\rangle_{\mathcal{M}([0,T];L^2(\Omega))\times\mathcal{C}([0,T];L^2(\Omega))}\leqslant 0,\quad\forall z\in\Theta,\tag{7.59}$$

$$\int_{\Sigma}\big(\lambda h_u'(x,\bar{y},\bar{u})+\widetilde{p}\big)(u-\bar{u})\,\mathrm{d}x\geqslant 0,\quad\forall u\in U_{ad}\tag{7.60}$$

成立. 由 (7.59) 和 (H4) 得 $\mu|_{\Omega\times\{0\}}$, 从而 μ 可等同于 $\mathcal{M}_b(]0,T];L^2(\Omega))$ 中的一个测度. 进而由定理 7.27 知, 伴随系统 (7.55) 存在满足 $(p,p_t)\in L^\infty\big(0,T;H^1(\Omega)\big)\times L^\infty\big(0,T;L^2(\Omega)\big)$ 的唯一弱解 $p(\cdot)$. 由定理 7.28 的 Green 公式和最优条件 (7.58) 得

$$\begin{aligned}&\big\langle p+\bar{p},y_{tt}-\Delta y\big\rangle+\int_{\Sigma}\Big(\widetilde{p}-\frac{\partial p}{\partial\nu}\Big)y\,\mathrm{d}s\,\mathrm{d}t\\&+\int_{\Omega}\big(\widehat{p}-p_t(0)\big)y(0)\,\mathrm{d}x+\int_{\Omega}\big(\breve{p}+p(0)\big)y_t(0)\,\mathrm{d}x=0\end{aligned}$$

对任意 $y\in Y$ 成立. 因为映射 $y\mapsto\big(y_{tt}-\Delta y,y|_{\Sigma},y(0),y_t(0)\big)$ 为从 Y 到 $L^1\big(0,T;H^{-1}(\Omega)\big)\times L^2(\Sigma)\times L^2(\Omega)\times H^{-1}(\Omega)$ 的满射, 上述变分条件给出

$$\begin{aligned}&p=-\bar{p}\in L^\infty\big(0,T;H_0^1(\Omega)\big),\quad\frac{\partial p}{\partial\nu}=\widetilde{p}\in L^2(\Sigma),\\&p_t(0)=\widehat{p}\in L^2(\Omega),\quad p(0)=-\breve{p}\in H_0^1(\Omega).\end{aligned}$$

这样定理 7.29 的必要最优条件 (7.58)~(7.60) 就蕴涵定理 7.23 中所需的最优条件 (7.50)~(7.52). 最后注意到定理 7.23 中的规范条件 (7.53) 可化为定理 7.29 中的对应条件, 这就保证了 $\lambda=1$. 证毕. △

注 7.30(SNC 状态约束) 由上述证明可知, 假设 $\operatorname{int}\Theta\neq\varnothing$ 可以本质地放宽为凸状态集 Θ 的 SNC 性质; 比照定理 5.11(ii) 及定理 2.51(ii) 的证明. 它们表明, 在一般的 Banach 空间中, 通过极限过程证明非平凡性只需要序列 (而不是拓扑) 法紧性就可以了. 对内点性质的要求在 7.2 节讨论的 Neumann 问题中也可以同样地放宽. 注意在 Banach 空间中凸集的情形, 若集合为有穷余维且有非空相对内部, 则 SNC 自动成立; 见定理 1.21. 另外, SNC 性质一般来说比上面的要求要弱, 在凸的情形事实上等价于 CEL(拓扑的) 性质; 见注 1.27 和例 3.6. 最后注意到, 上面证明中定理 5.11(ii) 的应用使得可以放宽 Dirchlet 边界控制问题 (DP) 中的被积项的可微性假设.

7.4 逐点状态约束下抛物系统的极小极大控制

本章最后一节讨论具有 Dirichlet 边界条件在逐点状态约束下的抛物控制系统. 把焦点放在 Dirichlet 边界控制问题上有以下两个原因. 第一, 抛物系统的 Dirichlet 情形比 Neumann 情形更复杂和具有挑战性, 这本质上不同于双曲的情形, 该情形在 Neumann 问题时更难但具有更多的正则性; 见前两节及相关的评注. 第二, 作者研究抛物系统控制问题的初始兴趣主要在于一些环境问题上的实际应用, 它们相关于土壤水的自动调节的变化规律, 见 Mordukhovich [898, 905]. 这些实际问题中的物理现象和工程建设导致了一些涉及具有 Dirichlet 边界控制的抛物方程模型. 进一步, 提到的这些系统的控制过程的实施都不可避免地碰到不确定的扰动, 最自然的优化判据就是极小极大. 考虑到这些, 本节考虑线性抛物系统的极小极大最优控制, 其控制作用于 Dirichlet 边界条件上, 具有不确定的分布扰动, 且状态和控制函数具有硬/逐点约束. 主要目标是建立极小极大解的存在性和最差扰动下开环控制的必要最优 (和次最优) 条件. 本节最后简略讨论了 (并给读者指出了进一步阅读的文献) 关于闭环抛物控制系统极小极大设计的一些课题; 涉及的 Dirichlet 边界条件的反馈控制已经超出了本书的范围而没有收录.

所论的极小极大问题本质上是非光滑的, 其变分分析需要特殊的方法. 为此本节系统地应用了光滑逼近过程. 事实上, 这里把原问题分拆成了对分布扰动和具有移动状态约束边界控制的两个相关的最优控制问题, 然后通过涉及极大单调算子的 C^∞-逼近的有效罚函数来逼近每个问题的状态约束, 并建立了这个过程的强收敛结果, 得到了逼近问题最优解的刻画. 最后给出恰当的状态规范条件, 得到原状态约束极小极大问题对最差扰动和控制的必要最优条件.

上述变分分析最复杂的部分在于涉及一个最差扰动下的状态约束 Dirichlet 边界控制问题. 主要难关是逐点状态约束和作用于 Dirichlet 边界条件上的 L^∞ 控制的硬约束同时出现的情形. 众所周知, 这些条件只能给出解的最少的正则性, 它们相关于变分不等式框架下的无界算子. 基于光滑逼近过程和抛物方程 Dirichlet 边界控制问题适度解的性质, 本节对此建立了有效的分析.

7.4.1 问题的表述与分拆

考虑如下抛物系统:

$$
\begin{cases}
\dfrac{\partial y}{\partial t} + Ay = Bw + \vartheta, & \text{在 } Q := (0,T)\times\Omega \text{ 中几乎处处成立,}\\
y(x,0) = y_0(x), & x\in\Omega,\\
y(s,t) = u(s,t), & (s,t)\in\Sigma := (0,T]\times\Gamma.
\end{cases}
\tag{7.61}
$$

它具有状态轨道 $y(\cdot)$ 上的逐点/硬约束、不确定扰动/干扰 $w(\cdot)$, 以及 Dirichlet 边界控制 $u(\cdot)$ 如下:

$$a \leqslant y(x,t) \leqslant b, \quad \text{a.e. } (x,t) \in Q, \tag{7.62}$$

$$c \leqslant w(x,t) \leqslant d, \quad \text{a.e. } (x,t) \in Q, \tag{7.63}$$

$$\mu \leqslant u(s,t) \leqslant \nu, \quad \text{a.e. } (s,t) \in \Sigma, \tag{7.64}$$

其中 $\Omega \subset \mathbb{R}^n$ 为具有足够光滑边界 Γ 的有界开集, 且每个区间 $[a,b]$, $[c,d]$, 和 $[\mu,\nu]$ 都包含 0.

令 $X := L^2(\Omega;\mathbb{R})$, $U := L^2(\Gamma;\mathbb{R})$ 和 $W := L^2(\Omega;\mathbb{R})$ 分别为状态、控制、干扰空间. 后面的叙述中不再在这些记号中写上 $\mathbb{R}$, 对类似的实值函数空间也类似处理. 记

$$U_{ad} := \left\{u \in L^p(0,T;U) \mid \mu \leqslant u(s,t) \leqslant \nu, \quad \text{a.e. } (s,t) \in \Sigma\right\}$$

为容许控制集, 其中 $L^p(0,T;U)$ 为 $[0,T]$ 上取值于 U 的函数 $u(\cdot) = u(s,\cdot)$ 的空间, 其范数定义为

$$\|u\|_{L^p(0,T;U)} := \left(\int_0^T \|u(t)\|_U^p \,\mathrm{d}t\right)^{1/p} = \left(\int_0^T \left(\int_\Gamma |u(s,t)|^2 \,\mathrm{d}s\right)^{p/2} \mathrm{d}t\right)^{1/p}.$$

类似地定义容许干扰集

$$W_{ad} := \left\{w \in L^2(0,T;W) \mid c \leqslant w(x,t) \leqslant d, \quad \text{a.e. } (x,t) \in Q\right\}.$$

二元组 $(u,w) \in U_{ad} \times W_{ad}$ 称为系统 (7.61) 的一个“可行解”, 如果对应的轨道 $y(\cdot)$ 满足状态约束 (7.62). 这里总是假设问题 (7.61)~(7.64) 至少具有一个可行二元组 (u,w).

尽管约束集 W_{ad} 和 U_{ad} 是本质有界的, 即 $W_{ad} \subset L^\infty(Q)$ 和 $U_{ad} \subset L^\infty(\Sigma)$, 本节倾向于把它们分别看成更大空间 $L^2(0,T;W)$ 和 $L^p(0,T;U)$ 的子集, 其中 p 有限但足够大, 原因是这些大空间具有自反性和范数在原点以外的可微性, 因而更容易有效地进行变分分析.

整个 7.4 节对所论的抛物系统有如下“常驻假设”:

(H1) 线性算子

$$A := -\sum_{i,j=1}^n \frac{\partial}{\partial x_i}\Big(a_{ij}(x)\frac{\partial}{\partial x_j}\Big) + \sum_{i=1}^n a_i(x)\frac{\partial}{\partial x_i} + a_0(x)$$

具有实值光滑系数且在 Ω 上“强一致椭圆”, 即存在 $\beta_0 > 0$ 满足

$$\sum_{i,j=1}^n a_{ij}(x)v_iv_j \geqslant \beta_0 \sum_{i=1}^n v_i^2, \quad \forall x \in \Omega, \quad (v_1,\cdots,v_n) \in \mathbb{R}^n.$$

(**H2**) $\vartheta \in L^{\infty}(Q)$, $y_0(x) \in H_0^1(\Omega) \cap H^2(\Omega)$, 且

$$a \leqslant y_0(x) \leqslant b, \quad \text{a.e. } x \in \Omega.$$

(**H3**) $B\colon L^2(0,T;W) \to L^2(0,T;X)$ 为有界线性算子.

下面的讨论中总是可以假设 $-A$ 生成 X 上的一个"强连续解析半群" $S(\cdot)$ 使得指数估计

$$\|S(t)\| \leqslant M_1 \mathrm{e}^{-\omega t}, \tag{7.65}$$

对某 $\omega > 0$ 和 $M_1 > 0$ 成立, 其中 $\|\cdot\|$ 表示从 X 到 X 算子的标准范数. 否则, 由标准的过程可以构造一个形式为 $\widetilde{A} := A + \widetilde{\omega} I$ 的稳定平移并具有上述性质.

注意到, 因为 $w \in L^2(0,T;W)$ 和 $u \in L^p(0,T;U)$, 抛物系统 (7.61) 对某些 $(u,w) \in U_{ad} \times W_{ad}$ 可能没有强或经典的解. 此时, 主要的困难来自 Dirichlet 边界条件中的不连续控制. 利用抛物方程的半群方法这个优势, 本节分析中使用 Dirichlet 边界问题的"适度解"这个概念.

考虑所谓的"Dirichlet 映射" D, 定义为 $y = Du$, 其中 $y(\cdot)$ 满足齐性椭圆方程

$$\begin{cases} -Ay = 0, & \text{在 } Q \text{ 中成立}, \\ y(s,t) = u(s,t), & (s,t) \in \Sigma. \end{cases}$$

众所周知 (比如见 Lions 与 Magenes [794]), Dirichlet 算子

$$D\colon L^2(\Gamma) \to \mathcal{D}(A^{1/4-\delta}) = H^{1/2-2\delta}(\Omega), \quad 0 < \delta \leqslant 1/4 \tag{7.66}$$

是连续线性的, 其中 $\mathcal{D}$ 按惯例表示定义域.

定义 7.31(Dirichlet 抛物系统的适度解) 连续映射 $y\colon [0,T] \to X$ 称为系统 (7.61) 对应于 $(u,w) \in L^p(0,T;U) \times L^2(0,T;W)$ 的"适度解", 如果对任意 $t \in [0,T]$ 成立如下表示:

$$\begin{aligned} y(t) = {} & S(t)y_0 + \int_0^t S(t-\tau)\big(Bw(\tau) + \vartheta(\tau)\big)\,\mathrm{d}\tau + \int_0^t AS(t-\tau)Du(\tau)\,\mathrm{d}\tau \\ & + S(t)y_0 + \int_0^t S(t-\tau)\big(Bw(\tau) + \vartheta(\tau)\big)\,\mathrm{d}\tau \\ & + \int_0^t A^{3/4+\delta} S(t-\tau) A^{1/4-\delta} Du(\tau)\,\mathrm{d}\tau, \end{aligned}$$

其中 D 为 (7.66) 中定义的对应于某个 $\delta \in (0,1/4]$ 的 Dirichlet 算子.

读者可以在 Lasiecka 与 Triggiani 的文章 [743] 及其文献中找到 Dirichlet 抛物系统适度解的更多信息. 特别地, 当 $p > 0$ 足够大时, 上面的假设保证了 (7.61) 的适度解对任意 $w \in L^2(0,T;W)$ 和 $u \in L^p(0,T;U)$ 都存在且唯一. 另外可以看到, 尽

管定义 7.31 中取值于 X 的函数 $y(t)$ 由定义是连续的, 二元实函数 $y(x,t)$ 却只是可测的, 这是为 $X = L^2(\Omega)$. 这就把抛物方程的适度解和其他解的概念明显地区别开来. 适度解允许人们处理涉及本节中所考虑的 Dirichlet 边界条件的抛物系统的非正则 (可测) 数据. 另一方面, 连续性的缺失造成了后面要克服的本质困难.

注意定义 7.31 中的 δ 可以是区间 $(0,1/4]$ 中的任意固定常数. 尽管该定义中 $y(t)$ 的第一个表示根本不依赖于 δ, 该常数却显式地出现在下述的一些估计中, 而且 δ 越接近于零则估计越好.

现在引入费用函数

$$J(u,w) := \int_Q g\big(x,t,y(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t + \int_Q f\big(x,t,w(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma h\big(s,t,u(s,t)\big)\,\mathrm{d}s\,\mathrm{d}t, \tag{7.67}$$

其中 $y(\cdot)$ 是系统 (7.61) 由 $u(\cdot)$ 和 $w(\cdot)$ 生成的轨道 (适度解). 这里总假设费用泛函 (7.67) 对问题 (7.61)~(7.64) 中的所有容许过程都有定义且有限. 在 7.4.2~7.4.4 小节中会给出被积项 g, f 和 h 的更多假设. 本节考虑的“极小极大控制问题”定义如下:

(P) 求容许控制 $\bar{u} \in U_{ad}$ 和干扰 $\bar{w} \in W_{ad}$ 使得 $(\bar{u},\bar{w})$ 为泛函 $J(u,w)$ 满足系统 (7.61) 和状态约束 (7.62) 的鞍点.

由鞍点的定义, 这意味着在条件 (7.61) 和 (7.62) 下有

$$J(\bar{u},w) \leqslant J(\bar{u},\bar{w}) \leqslant J(u,\bar{w}), \quad \forall u \in U_{ad}, \quad w \in W_{ad}. \tag{7.68}$$

这样的二元组 $(\bar{u},\bar{w})$ 称为极小极大问题 (P) 的最优解.

对问题 (P) 最优解的研究本节分为如下两个过程, 这显著地利用了系统 (7.61) 的线性性. 分拆的办法是, 把原问题 (7.61) 分成两个具有分离的干扰和边界控制的子系统. 第一个系统

$$\begin{cases} \dfrac{\partial y_1}{\partial t} + Ay_1 = Bw + \vartheta, & \text{在 } Q \text{ 中几乎处处成立}, \\ y_1(x,0) = y_0(x), & x \in \Omega, \\ y_1(s,t) = 0, & (s,t) \in \Sigma \end{cases} \tag{7.69}$$

具有零 (齐次) 边界条件并只依赖于干扰. 第二个系统

$$\begin{cases} \dfrac{\partial y_2}{\partial t} + Ay_2 = 0, & \text{在 } Q \text{ 中几乎处处成立}, \\ y_2(x,0) = 0, & x \in \Omega, \\ y_2(s,t) = u(s,t), & (s,t) \in \Sigma \end{cases} \tag{7.70}$$

由边界控制生成而不涉及干扰. 容易看到, 对任意 $(u,w) \in U_{ad} \times W_{ad}$, 关系

$$y(x,t) = y_1(x,t) + y_2(x,t) \quad (\forall (x,t) \in Q)$$

对系统 (7.61), (7.69) 和 (7.70) 的对应轨道成立.

令 $\bar{y}_1$ 和 $\bar{y}_2$ 分别为系统 (7.69) 和 (7.70) 对应于 $\bar{w}$ 和 $\bar{u}$ 的 (唯一) 轨道. 考虑对干扰 $w(\cdot)$ 的费用函数

$$J_1(w,y_1) := \int_Q \Big[g\big(x,t,y_1(x,t)+\bar{y}_2(x,t)\big)+f\big(x,t,w(x,t)\big)\Big]\,\mathrm{d}x\,\mathrm{d}t$$

和对边界控制 $u(\cdot)$ 的费用函数

$$J_2(u,y_2) := \int_Q g\big(x,t,\bar{y}_1(x,t)+y_2(x,t)\big)+\int_\Sigma h\big(s,t,u(s,t)\big)\,\mathrm{d}s\,\mathrm{d}t.$$

定义对应于这两个费用函数的两个优化问题如下. 第一个是

(P_1) 对 $w\in W_{ad}$ 极大化 $J_1(w,y_1)$ 使得 (7.69) 成立且

$$a-\bar{y}_2(x,t)\leqslant y_1(x,t)\leqslant b-\bar{y}_2(x,t),\quad \text{a.e. } (x,t)\in Q.$$

第二个是

(P_2) 对 $u\in U_{ad}$ 极小化 $J_2(u,y_2)$ 使得 (7.70) 成立且

$$a-\bar{y}_1(x,t)\leqslant y_2(x,t)\leqslant b-\bar{y}_1(x,t),\quad \text{a.e. } (x,t)\in Q.$$

下述结果表明, 原极小极大问题 (P) 可分拆成这两个状态约束动态优化问题 (P_1) 和 (P_2), 分别对应于干扰和控制.

命题 7.32(极小极大问题的分拆) 设 $(\bar{u},\bar{w})$ 为问题 (P) 的最优解, $\bar{y}_1$ 和 $\bar{y}_2$ 分别为系统 (7.69) 和 (7.70) 的对应轨道. 则 $\bar{w}$ 和 $\bar{u}$ 分别是 (P_1) 和 (P_2) 的解.

证明 由上述关系 $y(\cdot)=y_1(\cdot)+y_2(\cdot)$, 立刻得 $\bar{w}$ 为 (P_1) 的可行解, 即 (7.69) 对应的轨道 $\bar{y}_1$ 满足 (P_1) 中的状态约束. 于是由于费用函数 J 和 J_1 在所论问题中的构造, 则 (7.68) 的左边蕴涵 $\bar{w}$ 是 (P_1) 的最优解. 对 $\bar{u}$ 的证明类似. △

因此为得到极小极大问题 (P) 最优解 $(\bar{u},\bar{w})$ 的必要条件, 下面考虑两个分开的问题, 即关于 $\bar{w}$ 的 (P_1) 和关于 $\bar{u}$ 的 (P_2), 两个问题具有互相关联的状态约束. 这些约束依赖于 (x,t), 即是移动的. 对具有不确定扰动和作用于 Dirichlet 边界条件的非正则边界控制的原极小极大控制问题的研究来说, 这个性质是至关重要的.

7.4.2 适度解的性质和极小极大存在定理

本小节建立后面需要的抛物系统 (7.61) 适度解重要的正则性和收敛性, 然后证明极小极大最优解的存在性.

令 $S(t)$ 为 X 上由算子 $-A$ 生成的解析半群并满足指数估计 (7.65), 令 D 为具有连续性质 (7.66) 的 Dirichlet 算子. 下面用到估计

$$\|A^\delta D\|\leqslant M_2,\quad \|A^{3/4+\delta}S(t)\|\leqslant \frac{M_3}{t^{3/4+\delta}},\quad \forall\delta\in(0,1/4], \tag{7.71}$$

其中 $\|\cdot\|$ 表示对应的算子范数; 见 Lasiecka 和 Triggiani [743] 及其中的文献. 由定义 7.31 可以清楚地看到, 适度解研究的主要困难在于涉及 Dirichlet 算子的那一项. 下面利用从 $L^p(0,T;U)$ 到 $L^r\big(0,T;H^{1/2-\varepsilon}(\Omega)\big)$ 的算子 $\mathcal{L}$ 单独考虑这一项, 该算子定义为

$$\mathcal{L}u=(\mathcal{L}u)(t):=A\int_0^t S(t-\tau)Du(\tau)\,\mathrm{d}\tau$$
$$=\int_0^t A^{3/4+\delta}S(t-\tau)A^{1/4-\delta}Du(\tau)\,\mathrm{d}\tau, \tag{7.72}$$

其中 $p,r\in[1,\infty]$, $\delta\in(0,1/4]$, 且 $\varepsilon\in(0,1/2]$. 这里 Sobolev 空间 $H^{1/2-\varepsilon}(\Omega)\subset L^2(\Omega)=X$ 具有范数

$$\|y\|_{1/2-\varepsilon}:=\|A^{1/4-\varepsilon/2}y\|_X.$$

注意到 $H^0(\Omega)=L^2(\Omega)$ 且上面的范数 $\|y\|_{1/2-\varepsilon}$ 强于 $\|y\|_X$. 这里总在 (7.72) 中取 $\delta=\varepsilon/4$ 并称 $\mathcal{L}$ 为“适度解算子”. 该算子一般是无界的, 然而当 p 足够大时, 却具有很好的正则性/连续性如下.

定理 7.33(抛物 Dirichlet 系统适度解的正则性) 设 $\varepsilon\in(0,1/2]$, $p>4/\varepsilon$. 则 $\mathcal{L}u\in\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 对任意 $u\in L^p(0,T;U)$ 成立. 进一步, 算子

$$\mathcal{L}\colon L^p(0,T;U)\to\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$$

线性连续.

证明 显然 $\mathcal{L}$ 线性. 下面证明 $\mathcal{L}$ 有界从而连续. 即要证

$$\|\mathcal{L}u\|_{\mathcal{C}([0,T];H^{1/2-\varepsilon}(\Omega))}\leqslant K\|u\|_{L^p(0,T;U)}$$

对某 $K>0$ 成立. 由 (7.71) 和 (7.72) 知, 当 $t\in[0,T]$ 时有

$$\begin{aligned}\|(\mathcal{L}u)(t)\|_{1/2-\varepsilon}&=\Big\|\int_0^t A^{1/4-\varepsilon/2}AS(t-\tau)Du(\tau)\,\mathrm{d}\tau\Big\|_X\\&=\Big\|\int_0^t A^{1-\varepsilon/4}S(t-\tau)A^{1/4-\varepsilon/4}Du(\tau)\,\mathrm{d}\tau\Big\|_X\\&\leqslant M_2M_3\int_0^t(t-\tau)^{-(1-\varepsilon/4)}\|u\|_U\,\mathrm{d}\tau\\&\leqslant M_2M_3\Big(\int_0^t(t-\tau)^{-(1-\varepsilon/4)q}\,\mathrm{d}\tau\Big)^{1/q}\|u\|_{L^p(0,T;U)},\end{aligned}$$

其中 $1/p+1/q=1$. 因为 $p>4/\varepsilon$ 蕴涵 $q<4/(4-\varepsilon)$, 所以

$$\|(\mathcal{L}u)(t)\|_{1/2-\varepsilon}\leqslant M_2M_3\Big(\frac{1}{1-(1-\varepsilon/4)q}\Big)^{1/q}t^{(1-(1-\varepsilon/4)q)/q}\|u\|_{L^p(0,T;U)}.$$

下证 $\mathcal{L}u \in \mathcal{C}([0,T];H^{1/2-\varepsilon}(\Omega))$, 即 $(\mathcal{L}u)(t)$ 在 $H^{1/2-\varepsilon}(\Omega)$ 的范数下在任意点 $t_0 \in [0,T]$ 连续. 事实上, 确定起见, 取 $t \geqslant t_0$, 则

$$\begin{aligned}(\mathcal{L}u)(t)-(\mathcal{L}u)(t_0)=&\int_{t_0}^{t} AS(t-\tau)Du(\tau)\,\mathrm{d}\tau\\&+\big(S(t-t_0)-I\big)\int_{0}^{t_0} AS(t-\tau)Du(t)\,\mathrm{d}\tau.\end{aligned}$$

由上面的 $\|(\mathcal{L}u)(t)\|_{1/2-\varepsilon}$ 的估计和 $S(\cdot)$ 的强连续性, 有

$$\big\|(\mathcal{L}u)(t)-(\mathcal{L}u)(t_0)\big\|_{1/2-\varepsilon}\to 0 \quad (t\to t_0).$$

进一步, 由此估计和 $\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 中范数的定义, 立刻得所需的界限估计

$$K := M_2M_3\Big(\frac{1}{1-(1-\varepsilon/4)q}\Big)^{1/q}T^{(1-(1-\varepsilon/4)q)/q}.$$

证毕. △

推论 7.34(解算子的弱连续性) 设 ε 和 p 如定理 7.33 中所取. 则适度解算子 $\mathcal{L}$ 作为 $L^p(0,T;U)$ 到 $\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 的算子时是弱连续的. 这蕴涵对 $L^p(0,T;U)$ 中任意弱收敛序列 $u_k\to u$, 在 $k\to\infty$ 时沿 $\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 的弱拓扑有 $\mathcal{L}u_k\to\mathcal{L}u$.

证明 根据赋范空间之间线性连续算子弱连续性的标准事实, 由定理 7.33 即得. △

如前所述, (7.72) 中的算子 $\mathcal{L}$ 对定义 7.31 中适度解的结构至关重要; 这是它被称为适度解算子的原因. 由上述结果, 若 p 足够大, 则 $L^p(0,T;U)$ 中边界控制的强 (对应地, 弱) 收敛蕴涵系统 (7.61) 在空间 $\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 中对应轨道的强 (对应地, 弱) 收敛性. 可以看到, 如果定义中没有含 $\mathcal{L}$ 的这一项, 即 (7.61) 中 $u=0$ 的情形, 则 (7.61) 的适度解退化成通常意义下的标准 (强) 解. 特别地, 此时 $L^p(0,T;W)$ 中干扰的弱收敛 $w_k\to w$ 蕴涵对应的轨道在 $\mathcal{C}\big([0,T];X\big)$ 中的强收敛 $y_k\to y$ $(k\to\infty)$ 对任意 $p\geqslant 1$ 成立.

原问题 (P) 及其分拆问题的一个特殊性质是, 所有的约束都几乎处处由不连续的实二元 (空间和时间) 函数以逐点/硬的形式给出. 与此同时, 用于此处的半群方法涉及的是取值于泛函空间的连续依赖于时间的映射. 对后面的研究, 需要建立一个恰当的算子收敛性, 从而能蕴涵状态轨道的几乎处处收敛性. 下面这个结果对此方向是举足轻重的, 满足了进一步的变分分析的需求.

定理 7.35(适度解的逐点收敛性) 设 ε 和 p 如定理 7.33 中所取. 则 Dirichlet 控制在 $L^p(0,T;U)$ 中的弱收敛 $u_k\to u$ 蕴涵原抛物系统适度解算子 $\mathcal{L}$ 值的强收敛, 即在 $k\to\infty$ 时沿 $L^2(\Omega)$ 的强拓扑有

$$\mathcal{L}u_k \to \mathcal{L}u.$$

进一步, 存在 $\{(\mathcal{L}u_k)(x,t)\}$ 的实值子序列当 $k \to \infty$ 时在 Q 中几乎处处收敛到 $(\mathcal{L}u)(x,t)$.

证明　由推论 7.34 的弱收敛结果得, 对任意 $t \in [0,T]$ 在 $H^{1/2-\varepsilon}(\Omega)$ 的弱拓下有

$$(\mathcal{L}u_k)(\cdot,t) \to (\mathcal{L}u)(\cdot,t)$$

且序列 $\{\mathcal{L}u_k\}$ 在 $\mathcal{C}\big([0,T];H^{1/2-\varepsilon}(\Omega)\big)$ 中有界. 另外, 经典的嵌入结果保证了由 $H^{1/2-\varepsilon}(\Omega)$ 到 X 的嵌入映射的紧性; 例如见 Lions 与 Magenes [794, 定理 16.1]. 这蕴涵 X 中的强收敛

$$(\mathcal{L}u_k)(t,\cdot) \to (\mathcal{L}u)(t,\cdot) \quad \forall t \in [0,T].$$

因此有下述结论:

(i) 序列 $\{(\mathcal{L}u_k)(t,\cdot)\}$ 在 X 中有界, 即存在 $M \geqslant 0$ 满足估计

$$\big\|(\mathcal{L}u_k)(t)\big\|_X \leqslant M, \quad \forall t \in [0,T], \quad k \in \mathbb{N}.$$

(ii) 下述强收敛成立:

$$\big\|(\mathcal{L}u_k)(t) - (\mathcal{L}u)(t)\big\|_X \to 0, \quad \forall t \in [0,T] \quad (k \to \infty).$$

现在考虑由下述积分定义的 $[0,T]$ 上的实值非负函数序列 φ_k:

$$\varphi_k(t) := \int_\Omega \big|(\mathcal{L}u_k)(x,t) - (\mathcal{L}u)(x,t)\big|^2 \,\mathrm{d}x, \quad \forall t \in [0,T].$$

则 (i) 和 (ii) 分别蕴涵 φ_k 在 $[0,T]$ 上一致有界和 $\varphi_k(t) \to 0$ 在 $[0,T]$ 中逐点成立 $(k \to \infty)$. 由 Lebesgue 控制收敛定理得

$$\int_0^T \varphi_k(t)\,\mathrm{d}t \to 0 \quad (k \to \infty),$$

这保证了定理中的强算子收敛性. 进而得 $\{(\mathcal{L}u_k)(x,t)\}$ 包含一个对几乎处处的 $(x,t) \in Q$ 收敛到 $(\mathcal{L}u)(x,t)$ 的子序列. △

上面得到的这个收敛性/连续性结果对后面问题 (P), (P_1), (P_2) 的变分分析是基本的, 因为这样的变分分析大量涉及各种逼近过程的极限. 下面总是假设 (而不再具体说明) p 足够大以保证定理 7.35 的收敛结果成立.

接下来给出对费用函数 (7.67) 中被积项的假设, 它们保证了在对应的弱拓扑下该积分泛函关于变量 u 和 w 的恰当上半和下半连续性.

(**H4a**) $g(x,t,y)$ 满足 Carathéodory条件, 即其对 $(x,t)\in Q$ 可测 (对任意 $y\in\mathbb{R}$ 成立), 并对 $y\in\mathbb{R}$ 连续 (对几乎处处的 $(x,t)\in Q$ 成立). 另外, 存在非负函数 $\eta(\cdot)\in L^1(Q)$ 和常数 $\zeta\geqslant 0$ 满足

$$|g(x,t,y)|\leqslant \eta(x,t)+\zeta|y|^2,\quad \text{a.e. }(x,t)\in Q\quad (\text{当 } y\in\mathbb{R}).$$

(**H5a**) $f(x,t,w)$ 对 $(x,t)\in Q$ 可测, 对 $w\in[c,d]$ 连续凸, 并对某函数 $\kappa(\cdot)\in L^1(Q)$ 成立

$$f(x,t,w)\leqslant \kappa(x,t),\quad \text{a.e. }(x,t)\in Q\quad (\text{当 } w\in[c,d]).$$

(**H6a**) $h(s,t,u)$ 对 $(s,t)\in\Sigma$ 可测, 对 $u\in[\mu,\nu]$ 连续凸, 且对某函数 $\gamma(\cdot)\in L^1(\Sigma)$ 成立

$$h(s,t,u)\geqslant \gamma(s,t),\quad \text{a.e. }(s,t)\in\Sigma\quad (\text{当 } u\in[\mu,\nu]).$$

现在建立所论抛物系统极小极大最优解的存在性定理.

定理 7.36(极小极大解的存在性) 假设 (H1)~(H3) 和 (H4a)~(H6a) 成立, 并额外假设积分项 g 对于 y 线性. 则 (7.67) 中的费用函数 $J(u,w)$ 在满足抛物系统 (7.61) 的情况下具有在 $U_{ad}\times W_{ad}$ 中的鞍点 $(\bar{u},\bar{w})$. 进一步, 如果 (7.61) 的对应轨道满足状态约束 (7.62), 则 $(\bar{u},\bar{w})$ 是原极小极大问题 (P) 的最优解.

证明 考虑定义于集合 $U_{ad}\times W_{ad}\subset L^p(0,T;U)\times L^2(0,T;W)$ 上的泛函 $J(u,w)$, 其中 p 足够大. 可以看到, 集合 U_{ad} 和 W_{ad} 都是凸的, 且分别在自反空间 $L^p(0,T;U)$ 和 $L^2(0,T;W)$ 中弱列紧. 进一步根据假设易知, J 在 $U_{ad}\times W_{ad}$ 上是凸–凹的, 这里 g 对于 y 的线性性很关键.

现在验证恰当的半连续性以便在所论的无穷维空间中应用经典的 von Neumann 鞍点定理 (比如可见 Simons [1213]), 此时序列弱收敛和拓扑弱收敛是等价的. 首先证明, 对任意固定的 $w\in L^2(0,T;W)$, J 对于变量 u 在空间 $L^p(0,T;U)$ 中是序列弱下半连续的. 为此任取序列 $\{u_k\}$ 在空间 $L^p(0,T;U)$ 中弱收敛到 $\widetilde{u}$ $(k\to\infty)$. 由 Mazur 定理找到一个由 u_k 的凸组合构成的序列在 $L^p(0,T;U)$ 中强收敛到 $\widetilde{u}$. 因为 $U=L^2(\Sigma)$, 该序列也在 $L^2(\Sigma)$ 中强收敛到 $\widetilde{u}$. 由基于 h 对于 u 的凸性的标准推理和 (H6a) 中其他假设可得

$$\int_\Sigma h\big(s,t,\widetilde{u}(s,t)\big)\leqslant \liminf_{k\to\infty}\int_\Sigma h\big(s,t,u_k(s,t)\big)\,\mathrm{d}s\,\mathrm{d}t.$$

进一步考虑系统 (7.61) 对任意固定的 w 由 u_k 和 $\widetilde{u}$ 分别生成的轨道 (适度解) y_k 和 $\widetilde{y}$. 则根据定理 7.35 得 $y_k\to\widetilde{y}$ 沿 $L^2(Q)$ 的强拓扑在 $k\to\infty$ 时成立. 为得到

$$\int_Q g\big(x,t,\widetilde{y}(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t=\lim_{k\to\infty}\int_Q g\big(x,t,y_k(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t,$$

应用 Polyak [1096, 定理 2], 若 g 满足 (H4a) 中的 Carathéodory条件, 则 (H4a) 中的增长条件对 $L^2(Q)$ 中的积分泛函

$$I(y) := \int_Q g(x,t,y)\,\mathrm{d}x\,\mathrm{d}t$$

的强连续性是必要充分的. 因此对任意固定的 w, 在所给的假设下 (7.67) 中的费用泛函 $J(\cdot,w)$ 在 U_{ad} 上是序列弱下半连续的.

对任意固定的 u, 为证明 $J(u,\cdot)$ 在 W_{ad} 上的弱上半连续性, 由相同 (对称) 的推理, 并考虑到在 $L^2(0,T;W)$ 中的弱收敛 $w_k \to \widetilde{w}$ 直接蕴涵 $\mathcal{C}([0,T];X)$ 中的对应轨道的强收敛 $y_k \to \widetilde{y}$ (见紧接着推论 7.34 后的讨论), 则得 (7.67) 中的费用函数 $J(u,w)$ 在弱紧凸集 $U_{ad} \subset L^p(0,T;U)$ 上对于变量 u 是凸和弱下半连续的, 且在弱紧凸集 $W_{ad} \subset L^2(0,T;W)$ 上对于变量 w 是凹和上半连续的. 现在由无穷维空间中经典的极小极大定理就可以导出 J 在 $U_{ad} \times W_{ad}$ 满足系统 (7.61) 的鞍点 $(\bar{u},\bar{w})$ 的存在性. 进一步得, 如果对应轨道 $\bar{y}$ 满足状态约束 (7.62), 则显然 $(\bar{u},\bar{w})$ 是原问题 (P) 的最优解. 证毕. △

注 7.37(线性条件的放宽) (7.67) 中被积项上的假设 (H4a)~(H6a) 在整节都是需要的, 并在后面的关于逼近的稳定性及其变分分析的主要结果中将起到至关重要的作用. 相反地, g 对于 y 的苛刻的线性性假设只出现在定理 7.36 中以保证鞍点的存在性. 事实上只是用来得出费用函数 $J(u,w)$ 为凸–凹的. 在“混合策略”的框架下考虑鞍点时该假设可以去掉, 这类似本章及前面章节建立的最优控制问题的松弛过程. 另外可以看到, 基于本小节得到的正则性结果, g 对于 y 的线性性对下面两种情形的最优控制 (不是极小极大) 问题解的存在性是不必要的, 一是 Dirichlet 边界控制, 这是最难的情形, 二是分布控制和 Neumann 边界条件中的控制, 这对抛物系统容易处理一些; 更多细节见 Mordukhovich 和 Zhang [979].

7.4.3 最差扰动的次最优条件

本小节讨论 7.4.1 小节表述的第一个子问题 (P_1). 这里把 (P_1) 作为一个分布控制位于该抛物方程右边的最优控制问题. 这样极小极大问题 (P) 的“最差扰动”$\bar{w}(\cdot)$ 恰好是所论的具有移动状态约束的分布最优控制问题 (P_1) 的最优解 (极大化费用函数 J_1 的意义下). 注意这些移动状态约束涉及非正则 (可测) 函数 $\bar{y}_2(x,t)$, 即 Dirichlet 边界控制问题 (7.70) 的适度解, 实质性的困难正是由此造成的. 这里首先建立一种逼近方法消除这种约束, 并证明逼近的恰当强收敛性. 然后提供逼近问题详细的变分分析, 推导出最差扰动的必要次最优条件. 能建立必要最优条件的极限过程在 7.4.5 小节建立.

为此考虑集值极大单调算子 $\alpha\colon \mathbb{R} \rightrightarrows \mathbb{R}$, 形式为

$$\alpha(r):=\begin{cases}[0,\infty), & \text{若 } r=b,\\ (-\infty,0], & \text{若 } r=a,\\ 0, & \text{若 } a<r<b,\\ \varnothing, & \text{若 } r<a \text{ 或 } r>b.\end{cases}$$

由经典的“Moreau-Yosida 逼近”然后再用 $\mathbb{R}$ 上的 $\mathcal{C}_0^\infty$- 光滑化过程, 可以构造 $\alpha(\cdot)$ 的一系列光滑单值逼近 $\alpha_\varepsilon\colon \mathbb{R}\to\mathbb{R}$. 下面的这种逼近的具体实现对这里的研究是方便的. 构造 $\alpha_\varepsilon\colon \mathbb{R}\to\mathbb{R}$ $(\varepsilon>0)$ 为

$$\alpha_\varepsilon(r):=\begin{cases}\varepsilon^{-1}(r-b)-1/2, & \text{若 } r\geqslant b+\varepsilon,\\ (2\varepsilon^2)^{-1}(r-b)^2, & \text{若 } b\leqslant r<b+\varepsilon,\\ \varepsilon^{-1}(r-a)+1/2, & \text{若 } r\leqslant a-\varepsilon,\\ -(2\varepsilon^2)^{-1}(r-a)^2, & \text{若 } a-\varepsilon<r\leqslant a,\\ 0, & \text{若 } a<r<b.\end{cases} \tag{7.73}$$

对 $\alpha_\varepsilon(\cdot)$ 的导数, 易计算验证

$$\varepsilon\alpha_\varepsilon'(r):=\begin{cases}1, & \text{若 } r\geqslant b+\varepsilon,\\ \varepsilon^{-1}(r-b), & \text{若 } b\leqslant r<b+\varepsilon,\\ 1, & \text{若 } r\leqslant a-\varepsilon,\\ -\varepsilon^{-1}(r-a), & \text{若 } a-\varepsilon<r\leqslant a,\\ 0, & \text{若 } a<r<b,\end{cases}$$

其中 $|\varepsilon\alpha_\varepsilon'(r)|\leqslant 1$ 对任意 $r\in\mathbb{R}$ 成立.

令 $(\bar u,\bar w)$ 为极小极大问题 (P) 的给定最优解, $\bar y_1$ 和 $\bar y_2$ 分别是对应的系统 (7.69) 和 (7.70) 的轨道. 考虑下面的一类以 ε 为参数的没有状态约束的控制问题, 它们逼近 7.4.1 中第一个子问题 (P_1), 依赖于 Dirichlet 边界控制问题 (7.70) 的给定轨道 $\bar y_2$:

$(\mathrm{P}_{1\varepsilon})$ 极大化罚函数

$$\begin{aligned}J_{1\varepsilon}(w,y_1):=&\int_Q\Big[g\big(x,t,y_1(x,t)+\bar y_2(x,t)\big)+f\big(x,t,w(x,t)\big)\Big]\,\mathrm{d}x\,\mathrm{d}t\\ &-\big\|w-\bar w\big\|^2_{L^2(0,T;W)}-\varepsilon\big\|\alpha_\varepsilon(y_1+\bar y_2)\big\|^2_{L^2(0,T;X)},\\ \text{s.t.}\quad & w\in W_{ad}\text{并满足系统(7.69)}.\end{aligned}$$

因为 $w\in W_{ad}$ 和 $\vartheta\in L^\infty(Q)$, 经典的结果保证了, 具有齐次 Dirichlet 边界条件的抛物系统 (7.69) 具有“唯一强解” $y_1\in W^{1,2}([0,T];X)$ 满足估计

$$\Big\|\frac{\partial y_1}{\partial t}\Big\|_{L^2(0,T;X)}+\big\|Ay_1\big\|_{L^2(0,T;X)}\leqslant C\Big(\|y_0\|_{H_0^1(\Omega)\cap H^2(\Omega)}+\|Bw+\vartheta\|_{L^2(0,T;X)}\Big).$$

取 $\{w_k\} \subset W_{ad}$ 及对应的系统 (7.69) 的强解序列 $\{y_{1k}\}$, 并应用标准的推理 (比照 7.4.2 小节) 可得, 如果沿 $L^2(0,T;W)$ 的弱拓扑有 $w_k \to w \in W_{ad}$, 则 $y_{1k} \to y_1$ 沿 $\mathcal{C}\big([0,T];X\big)$ 的强拓扑成立 $(k \to \infty)$, 且 y_1 是系统 (7.69) 对应于 w 的强解.

接下来研究逼近问题 $(\mathrm{P}_{1\varepsilon})$. 第一个目标是对任意 $\varepsilon > 0$ 证明 $(\mathrm{P}_{1\varepsilon})$ 最优解的存在性. 这可化归为经典的 Weierstrass 定理, 它保证了在恰当拓扑下紧集上的上半连续费用函数总存在全局极大点. 这里的主要困难在于费用函数中的扰动项, 它依赖于 Dirichlet 系统 (7.70) 的非正则适度解 $\bar{y}_2$. 下面是这个结果及其技术上很复杂的证明.

定理 7.38(分布扰动逼近问题最优解的存在性) 设初始状态 y_0 满足假设 (H2) 且 $\varepsilon > 0$. 则逼近问题 $(\mathrm{P}_{1\varepsilon})$ 至少有一个最优解 $(w_\varepsilon, y_{1\varepsilon}) \in W_{ad} \times W^{1,2}\big([0,T];X\big)$.

证明 注意到问题 $(\mathrm{P}_{1\varepsilon})$ 的可行解集非空, 这是因为 $(\bar{w}, \bar{y}_1)$ 对任意 $\varepsilon > 0$ 都是 $(\mathrm{P}_{1\varepsilon})$ 的可行解. 下证 $(\mathrm{P}_{1\varepsilon})$ 中的费用函数 $J_{1\varepsilon}$ 是正常且上一致有界的, 即

$$j_{1\varepsilon} := \sup_{w \in W_{ad}} J_{1\varepsilon}(w, y_1) < \infty, \tag{7.74}$$

其中 $y_1 \in W^{1,2}\big([0,T];X\big)$ 为系统 (7.69) 对应的强解. 事实上, (H4a) 和 (H5a) 立刻蕴涵

$$\int_Q g\big(x,t,y_1(x,t) + \bar{y}_2(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t + \int_Q f\big(x,t,w(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t$$

对于 $w \in W_{ad}$ 的上一致有界性. 进一步, 显然存在 $\gamma > 0$ 满足

$$\big\|w - \bar{w}\big\|_{L^2(0,T;W)} < \gamma \quad (\text{当} w \in W_{ad}).$$

剩下需要分析依赖于 $\bar{y}_2$ 的 $J_{1\varepsilon}$ 的最后一项. 由估计 (7.71) 和适度解的定义 7.31 得

$$\big\|\bar{y}_2(t)\big\|_X \leqslant \frac{4M_2M_3 \max\big\{|\mu|, \nu\big\}\sqrt{\mathrm{mes}\,(\varGamma)}}{1-4\delta} t^{\frac{1-4\delta}{4}} \quad (\text{当}\delta \in (0, 1/4)),$$

其中 "mes" 表示一个集合的标准 Lebesgue 测度. 为估计 $\big\|\alpha_\varepsilon(y_1 + \bar{y}_2)\big\|_{L^2(0,T;X)}$, 考虑集合

$$\begin{aligned}
\varOmega_{1a}^t &:= \big\{x \in \varOmega \mid a - \varepsilon < y_1(x,t) + \bar{y}_2(x,t) \leqslant a\big\},\\
\varOmega_{2a}^t &:= \big\{x \in \varOmega \mid y_1(x,t) + \bar{y}_2(x,t) \leqslant a - \varepsilon\big\},\\
\varOmega_{1b}^t &:= \big\{x \in \varOmega \mid b \leqslant y_1(x,t) + \bar{y}_2(x,t) < b + \varepsilon\big\},\\
\varOmega_{2b}^t &:= \big\{x \in \varOmega \mid y_1(x,t) + \bar{y}_2(x,t) \geqslant b + \varepsilon\big\}.
\end{aligned}$$

对几乎处处的 $t \in [0,T]$, 它们都是 $\varOmega$ 的 Lebesgue 可测子集. 计入 (7.73) 中逼近结构 $\alpha_\varepsilon(\cdot)$ 和对任意 $r, s \in \mathbb{R}$ 成立的平凡不等式 $2(r^2 + s^2) \geqslant (r+s)^2$, 则有估计

$$
\begin{aligned}
&\left\|\alpha_\varepsilon(y_1+\bar{y}_2)\right\|_{L^2(0,T;X)}\\
=&\left[\int_0^T\int_\Omega \alpha_\varepsilon^2\big(y_1(x,t)+\bar{y}_2(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
=&\left[\int_0^T\left(\int_{\Omega_{1a}^t}(2\varepsilon^2)^{-2}\big(y_1(x,t)+\bar{y}_2(x,t)-a\big)^4\,\mathrm{d}x\right.\right.\\
&+\int_{\Omega_{2a}^t}\left(\varepsilon^{-1}\big(y_1(x,t)+\bar{y}_2(x,t)-a\big)+\frac{1}{2}\right)^2\mathrm{d}x\\
&+\int_{\Omega_{1b}^t}(2\varepsilon^2)^{-2}\big(y_1(x,t)+\bar{y}_2(x,t)-b\big)^4\,\mathrm{d}x\\
&\left.\left.+\int_{\Omega_{2b}^t}\left(\varepsilon^{-1}\big(y_1(x,t)+\bar{y}_2(x,t)-b\big)-\frac{1}{2}\right)^2\mathrm{d}x\right)\mathrm{d}t\right]^{1/2}\\
\leqslant&\left[\int_0^T\left(\frac{1}{4}\mathrm{mes}\,(\Omega_{1a}^t)+\frac{1}{4}\mathrm{mes}\,(\Omega_{1b}^t)\right)\mathrm{d}t\right]^{1/2}\\
&+\left[\int_0^T\int_{\Omega_{2a}^t}\left(\varepsilon^{-1}\big(y_1(x,t)+\bar{y}_2(x,t)\big)-\varepsilon^{-1}a+\frac{1}{2}\right)^2\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
&+\left[\int_0^T\int_{\Omega_{2b}^t}\left(\varepsilon^{-1}\big(y_1(x,t)+\bar{y}_2(x,t)\big)-\varepsilon^{-1}b-\frac{1}{2}\right)^2\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
\leqslant&\frac{1}{2}\sqrt{\mathrm{mes}\,(Q)}+\left[\int_0^T\int_{\Omega_{2a}^t}\left(2\varepsilon^{-2}\big(y_1(x,t)+\bar{y}_2(x,t)\big)^2+2\left(\varepsilon^{-1}|a|+\frac{1}{2}\right)^2\right)\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
&+\left[\int_0^T\int_{\Omega_{2b}^t}\left(2\varepsilon^{-2}\big(y_1(x,t)+\bar{y}_2(x,t)\big)^2+2\left(\varepsilon^{-1}b+\frac{1}{2}\right)^2\right)\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
\leqslant&\frac{1}{2}\sqrt{\mathrm{mes}\,(Q)}+\left[2\varepsilon^{-2}\int_0^T\int_{\Omega_{2a}^t}\big(y_1(x,t)+\bar{y}_2(x,t)\big)^2\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
&+\left[2\left(\frac{1}{2}+\varepsilon^{-1}|a|\right)^2\int_0^T\int_{\Omega_{2a}^t}\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
&+\left[2\varepsilon^{-2}\int_0^T\int_{\Omega_{2b}^t}\big(y_1(x,t)+\bar{y}_2(x,t)\big)^2\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
&+\left[2\left(\frac{1}{2}+\varepsilon^{-1}b\right)^2\int_0^T\int_{\Omega_{2b}^t}\mathrm{d}x\,\mathrm{d}t\right]^{1/2}\\
\leqslant&\frac{1}{2}\sqrt{\mathrm{mes}\,(Q)}+\sqrt{2}\,\varepsilon^{-1}\left\|y_1+\bar{y}_2\right\|_{L^2(0,T;X)}+\sqrt{2}\left(\frac{1}{2}+\varepsilon^{-1}|a|\right)\sqrt{\mathrm{mes}\,(Q)}\\
&+\sqrt{2}\,\varepsilon^{-1}\left\|y_1+\bar{y}_2\right\|_{L^2(0,T;X)}+\sqrt{2}\left(\frac{1}{2}+\varepsilon^{-1}b\right)\sqrt{\mathrm{mes}\,(Q)}\\
=&\left(\frac{1}{2}+\sqrt{2}+\varepsilon^{-1}b\sqrt{2}+\varepsilon^{-1}|a|\sqrt{2}\right)\sqrt{\mathrm{mes}\,(Q)}+2\sqrt{2}\,\varepsilon^{-1}\left\|y_1+\bar{y}_2\right\|_{L^2(0,T;X)}.
\end{aligned}
$$

再结合上述对 $\|\bar{y}_2(t)\|_X$ 和 (7.69) 的强解 $y_1 \in W^{1,2}([0,T];X)$ 的估计, 即得 $\|\alpha_\varepsilon(y_1+\bar{y}_2)\|_{L^2(0,T;X)}$ 在 $\varepsilon>0$ 时的一致有界性, 从而有 (7.74).

接下来, 对任意 $\varepsilon>0$ (简化起见可以略掉) 有 $(\mathrm{P}_{1\varepsilon})$ 可行解组成的极大化序列 $\{w_k, y_{1k}\}$ 满足

$$j_{1\varepsilon}-\frac{1}{k} \leqslant J_{1\varepsilon}(w_k, y_{1k}) \leqslant j_{1\varepsilon}, \quad \forall k \in \mathbb{N}. \tag{7.75}$$

因为 W_{ad} 在 $L^2(0,T;W)$ 中有界闭凸, 可以找到 $\{w_k\}$ 的一个子序列 (不再重新标识下标) 在 $L^2(0,T;W)$ 中弱收敛到某个 $\widetilde{w} \in W_{ad}$. 取系统 (7.69) 对应的 (强) 解 $\widetilde{y}_1$ 并再用上述的讨论, 则在 $k\to\infty$ 时有 $\mathcal{C}([0,T];X)$ 中的强收敛 $y_{1k} \to \widetilde{y}_1$. 由 (H4a) 和 (H5a) 及函数 $-\|\cdot\|^2_{L^2(0,T;W)}$ 的连续性和凹性, 则有 (比照定理 7.36 的证明)

$$\begin{aligned}
&\limsup_{k\to\infty}\left(\int_Q \Big[g\big(x,t,y_{1k}(x,t)+\bar{y}_2(x,t)\big)+f\big(x,t,w_k(x,t)\big)\Big]\,\mathrm{d}x\,\mathrm{d}t\right.\\
&\left.-\|w_k-\bar{w}\|^2_{L^2(0,T;W)}\right)\\
\leqslant &\int_Q \Big[g\big(x,t,\widetilde{y}_1(x,t)+\bar{y}_2(x,t)\big)+f\big(x,t,\widetilde{w}(x,t)\big)\Big]\,\mathrm{d}x\,\mathrm{d}t-\|\widetilde{w}-\bar{w}\|^2_{L^2(0,T;W)}.
\end{aligned}$$

因此

$$\lim_{k\to\infty}\alpha^2_\varepsilon(y_{1k}+\bar{y}_2)=\alpha^2_\varepsilon(\widetilde{y}_1+\bar{y}_2)$$

在 Q 中几乎处处成立. 在 $k\to\infty$ 时在 (7.75) 中取极限即得 $j_{1\varepsilon}=J_{1\varepsilon}(\widetilde{w},\widetilde{y}_1)$, 从而完成了证明. △

下面这个技术性的引理对证明本节逼近过程中状态约束的保持是重要的.

引理 7.39(状态约束的保持) 设 $y_k(x,t)(k\in\mathbb{N})$ 和 $y(x,t)$ 为 $L^1(Q)$ 中的非负函数. 给定 $c\geqslant 0$, 考虑集合

$$Q_k := \left\{(x,t)\in Q \mid y_k(x,t) > c+\frac{1}{k}\right\}, \quad k\in\mathbb{N}.$$

设 $y_k(x,t)\to y(x,t)$ 在 Q 中几乎处处成立, 且

$$\int_{Q_k} y_k(x,t)\,\mathrm{d}x\,\mathrm{d}t \to 0 \quad (k\to\infty),$$

则状态约束不等式 $0\leqslant y(x,t)\leqslant c$ 在 Q 中几乎处处成立.

证明 用反证法. 设结论不真. 则对任意足够小的 $\rho>0$, 存在可测集 $Q_\rho\subset Q$, 使得 $\mathrm{mes}\,(Q_\rho)>0$ 及

$$y(x,t)>c+\rho \quad (当(x,t)\in Q_\rho). \tag{7.76}$$

因为 $y_k(x,t) \to y(x,t)$ 在 Q 中几乎处处成立, 并应用实函数理论中经典的 Egorov 定理, 对任意 $\varepsilon > 0$ 和 $\rho > 0$, 存在可测集 $Q_\varepsilon \subset Q$ 和自然数 $k_\varepsilon \in \mathbb{N}$ (皆与 (x,t) 无关) 使得

$$\rho - \frac{1}{k} > \frac{\rho}{2} > 0, \quad \text{mes}\,(Q \backslash Q_\varepsilon) < \varepsilon,$$

$$|y_k(x,t) - y(x,t)| < \frac{\rho}{2} < \rho - \frac{1}{k} \quad (\text{当} k \geqslant k_\varepsilon, (x,t) \in Q_\varepsilon).$$

选取 $\varepsilon > 0$ 满足 $\text{mes}\,(Q_\rho \cap Q_\varepsilon) \neq 0$. 由 (7.76) 得

$$y_k(x,t) > y(x,t) - \rho + \frac{1}{k} > c + \rho - \rho + \frac{1}{k} = c + \frac{1}{k}$$

$$\big(\text{当} k > k_\varepsilon, \quad (x,t) \in Q_\rho \cap Q_\varepsilon\big).$$

这给出 $(Q_\rho \cap Q_\varepsilon) \subset Q_k \ (\forall k > k_\varepsilon)$. 因此由引理中的收敛性假设得

$$\int_{Q_\rho \cap Q_\varepsilon} y_k(x,t)\,\mathrm{d}x\,\mathrm{d}t \to 0 \quad (k \to \infty).$$

由于有 $y_k(x,t) \to y(x,t)$ 在 $Q_\rho \cap Q_\varepsilon$ 中一致成立 ($k \to \infty$), 这蕴涵条件

$$\int_{Q_\rho \cap Q_\varepsilon} y(x,t)\,\mathrm{d}x\,\mathrm{d}t = 0.$$

因为 $y(x,t) \geqslant 0$, 即有 $y(x,t) = 0$ 在 Q 中几乎处处成立, 这与 (7.76) 矛盾. 证毕. △

下面的定理建立了本小节发展的逼近过程的强收敛性. 这就说明, 逼近问题 $(\mathrm{P}_{1\varepsilon})$ 的最优解 (根据定理 7.38, 最优解一定存在) 恰好是状态约束问题 (P_1) 的次最优解, 而该问题对应于原来极小极大问题的最差扰动.

定理 7.40(最差扰动逼近问题的强收敛) 设 $(\bar{w}, \bar{y}_1)$ 为问题 (P_1) 的给定最优解, $\{(w_\varepsilon, y_{1\varepsilon})\}$ 为问题 $(\mathrm{P}_{1\varepsilon})$ 的最优解. 则存在 ε 的子序列, 当 $\varepsilon \downarrow 0$ 时, 有

$$w_\varepsilon \to \bar{w} \quad (L^2(0,T;W) \text{ 中的强收敛}),$$

$$y_{1\varepsilon} \to \bar{y}_1 \quad \big(\mathcal{C}([0,T];X) \text{ 中的强收敛}\big),$$

$$J_{1\varepsilon}(w_\varepsilon, y_{1\varepsilon}) \to J_1(\bar{w}, \bar{y}_1).$$

证明 应用定理 7.38 证明中的弱紧性推理, 可找到函数 $\widetilde{w} \in W_{ad}$ 和 $\{w_\varepsilon\}$ 的一个子序列使得

$$w_\varepsilon \to \widetilde{w} \quad (L^2(0,T;W) \text{ 中的强收敛}, \varepsilon \downarrow 0).$$

如前所证, 存在由 $\widetilde{w}$ 生成的系统 (7.69) 的 (唯一) 强解 $\widetilde{y}_1 \in W^{1,2}([0,T];X)$ 使得

$$y_{1\varepsilon} \to \widetilde{y}_1 \quad \big(\mathcal{C}([0,T];X) \text{ 中的强收敛}, \varepsilon \downarrow 0\big).$$

需要证明二元组 $(\widetilde{w},\widetilde{y}_1)$ 为问题 (P_1) 的可行解. 其实剩下只要证明 $\widetilde{y}$ 满足状态约束 (7,62) 就可以了, 即

$$a \leqslant \widetilde{y}_1(x,t)+\bar{y}_2(x,t) \leqslant b$$

在 Q 中几乎处处成立. 为此, 首先注意到 (P_1) 的最优解 $(\bar{w},\bar{y}_1)$ 对 $(\mathrm{P}_{1\varepsilon})$ 是可行的, 且 $\alpha_\varepsilon(\bar{y}_1+\bar{y}_2)=0$ 在 Q 中几乎处处成立 $(\forall\varepsilon>0)$. 因为 $(w_\varepsilon,y_{1\varepsilon})$ 是 $(\mathrm{P}_{1\varepsilon})$ 的最优解, 有

$$J_1(\bar{w},\bar{y}_1)=J_{1\varepsilon}(\bar{w},\bar{y}_1)\leqslant J_{1\varepsilon}(w_\varepsilon,y_{1\varepsilon}),\quad \forall\varepsilon>0. \tag{7.77}$$

应用此式并计入问题 $(\mathrm{P}_{1\varepsilon})$ 中费用函数的结构以及假设 (H4a) 和 (H5a), 则得 $\{\varepsilon^{1/2}\|\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\|_{L^2(0,T;X)}\}$ 有界. 于是

$$\varepsilon\big\|\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\big\|_{L^2(0,T;X)}\to 0\quad(\varepsilon\downarrow 0).$$

根据上述 $\alpha_\varepsilon(\cdot)$ 的构造和 Ω 的分划, 这给出

$$\begin{aligned}
&\int_0^T\int_{\Omega_{1a}^t}(2\varepsilon)^{-2}\Big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-a\Big)^4\mathrm{d}x\,\mathrm{d}t\\
&+\int_0^T\int_{\Omega_{2a}^t}\Big(\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-a\big)+\frac{\varepsilon}{2}\Big)^2\mathrm{d}x\,\mathrm{d}t\\
&+\int_0^T\int_{\Omega_{1b}^t}(2\varepsilon)^{-2}\Big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-b\Big)^4\mathrm{d}x\,\mathrm{d}t\\
&+\int_0^T\int_{\Omega_{2b}^t}\Big(\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-b\big)-\frac{\varepsilon}{2}\Big)^2\mathrm{d}x\,\mathrm{d}t\to 0\quad(\varepsilon\downarrow 0).
\end{aligned}$$

注意到对 $t\in[0,T]$ 几乎处处有

$$\begin{aligned}
&\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-a\big)^4\leqslant\varepsilon^4\quad(\text{在 }\Omega_{1a}^t\text{ 中几乎处处成立}),\\
&\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)-b\big)^4\leqslant\varepsilon^4\quad(\text{在 }\Omega_{1b}^t\text{ 中几乎处处成立})
\end{aligned}$$

并应用引理 7.39, 得极限 $(\widetilde{w},\widetilde{y}_1)$ 满足状态约束 (7.62), 从而是 (P_1) 的可行解. 因而 $J_1(\widetilde{w},\widetilde{y}_1)\leqslant J_1(\bar{w},\bar{y}_1)$.

现在由此事实证明定理中的强收敛性. 首先重写 (7.77) 为

$$J_1(\bar{w},\bar{y}_1)+\varepsilon\big\|\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\big\|^2_{L^2(0,T;X)}+\big\|w_\varepsilon-\bar{w}\big\|^2_{L^2(0,T;W)}\leqslant J_1(w_\varepsilon,y_{1\varepsilon})$$

并在不等式两边取上极限. 注意泛函 $J_1(w,y)$ 在所给的假设下对于 $L^2(0,T;W)$ 的弱拓扑和 $\mathcal{C}([0,T];X)$ 的范数拓扑都是上半连续的 (比照定理 7.38 的证明). 由此并

应用上面建立的弱收敛 $w_\varepsilon \to \widetilde{w}$ 和强收敛 $y_{1\varepsilon} \to \widetilde{y}_1$, 则得

$$
\begin{aligned}
&J_1(\bar{w}, \bar{y}_1) + \limsup_{\varepsilon\downarrow 0} \left(\varepsilon \left\|\alpha_\varepsilon(y_{1\varepsilon} + \bar{y}_2)\right\|^2_{L^2(0,T;X)} + \left\|w_\varepsilon - \bar{w}\right\|^2_{L^2(0,T;W)}\right) \\
\leqslant &\limsup_{\varepsilon\downarrow 0} J_1(w_\varepsilon, y_{1\varepsilon}) \leqslant J_1(\widetilde{w}, \widetilde{y}_1) \leqslant J_1(\bar{w}, \bar{y}_1).
\end{aligned}
$$

这蕴涵

$$
\lim_{\varepsilon\downarrow 0} \varepsilon \left\|\alpha_\varepsilon(y_{1\varepsilon} + \bar{y}_2)\right\|^2_{L^2(0,T;X)} = 0, \quad \lim_{\varepsilon\downarrow 0} \left\|w_\varepsilon - \bar{w}\right\|^2_{L^2(0,T;W)} = 0,
$$

即 w_ε 在 $L^2(0,T;W)$ 中强收敛到 $\bar{w}$, 从而 $y_{1\varepsilon}$ 在 $\mathcal{C}([0,T];X)$ 中强收敛到 $\bar{y}_1$, 其中 $\varepsilon \downarrow 0$. 最后, 定理中的值收敛可由 $J_1(\cdot)$ 在 $L^2(Q)$ 中强拓扑下的连续性 (见定理 7.36 证明中的讨论) 导出. 证毕. △

本小节的最后推导出对任意 $\varepsilon > 0$ 逼近问题 $(P_{1\varepsilon})$ 的必要最优条件. 根据定理 7.40 和分拆过程, 得到的结果可以看成原来极小极大问题中最差扰动的次最优条件. 问题 (P_1) 的必要最优条件将在 7.4.5 小节建立 (由定理 7.40 并在 $\varepsilon \downarrow 0$ 时对问题 $(P_{1\varepsilon})$ 的对应条件取极限).

考虑到容许扰动集 W_{ad} (现在是所论问题 $(P_{1\varepsilon})$ 的控制集合) 的凸性并且 $(P_{1\varepsilon})$ 没有状态约束, 这里用经典的控制变分和 7.4.2 小节中的正则性结果对每个 $(P_{1\varepsilon})$ 进行变分分析. 简化起见, 被积项对于控制/扰动和状态变量都加了某些光滑性. 如 6.4 节和 7.2 节, 如果涉及针形变分, 则可以减弱所给的光滑性和凸性的假设, 这里不再赘述. 现给出如下假设:

(H4b) 对几乎处处的 $(x,t) \in Q$, $g(x,t,y)$ 对 y 是连续可微的, 并且对任意 $y \in \mathbb{R}$, $(\partial g/\partial y)(x,t,y)$ 对于 (x,t) 是可测的. 进一步, 存在非负函数 $\eta_1 \in L^2(Q)$ 和常数 $\zeta_1 \geqslant 0$ 满足

$$
\left|\frac{\partial g}{\partial y}(x,t,y)\right| \leqslant \eta_1(x,t) + \zeta_1|y|
$$

对 $(x,t) \in Q$ 几乎处处成立, 其中 $y \in \mathbb{R}$ 是任意的.

(H5b) 对几乎处处的 $(x,t) \in Q$, $f(x,t,w)$ 对 w 是连续可微的, 且对任意 $w \in [c,d]$, $(\partial f/\partial w)(x,t,w)$ 对于 (x,t) 是可测的. 进一步, 存在非负函数 $\kappa_1 \in L^1(Q)$ 满足

$$
\left|\frac{\partial f}{\partial w}(x,t,w)\right| \leqslant \kappa_1(x,t)
$$

对 $(x,t) \in Q$ 几乎处处成立, 其中 $w \in [c,d]$ 是任意的.

考虑具有齐次 Dirichlet 边界条件的伴随抛物系统

$$\begin{cases} \dfrac{\partial p_1}{\partial t} - A^* p_1 = -\dfrac{\partial g}{\partial y}(x,t,y_{1\varepsilon} + \bar{y}_2) \\ \qquad\qquad + 2\varepsilon \alpha_\varepsilon'(y_{1\varepsilon} + \bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon} + \bar{y}_2) \quad \text{在 } Q \text{ 中几乎处处成立}, \\ p_1(T,x) = 0, \quad x \in \mathrm{cl}\,\Omega, \\ p_1(s,t) = 0, \quad (s,t) \in \Sigma, \end{cases} \tag{7.78}$$

其中 $\mathrm{cl}\,\Omega = \Omega \cup \Gamma$. 由 (H4b) 得

$$\frac{\partial g}{\partial y}\big(x,t,y_1(x,t) + \bar{y}_2(x,t)\big) \in L^2(\Omega) \quad \big(\text{当} y_1 \in \mathcal{C}\big([0,T];X\big)\big).$$

在经典的抛物方程理论中人们熟知, (7.78) 具有唯一强解 $p_{1\varepsilon} \in W^{1,2}\big([0,T];X\big)$ 满足

$$p_{1\varepsilon} \in \mathcal{C}\big([0,T];X\big) \cap L^2\big(0,T;H_0^1(\Omega) \cap H^2(\Omega)\big).$$

下述定理给出了逼近问题 $(\mathrm{P}_{1\varepsilon})$ 的必要最优条件, 它具有 (线性化了的) 最大值原理的积分形式. 由 W_{ad} 的约束构造, 很容易就蕴涵对应的砰–砰形式中的逐点结果; 见后面的推论. 下面逼近参数 $\varepsilon > 0$ 是固定的.

定理 7.41(最差扰动积分形式的次最优条件)　设 $(w_\varepsilon, y_{1\varepsilon})$ 为问题 $(\mathrm{P}_{1\varepsilon})$ 的一个最优解, $p_{1\varepsilon}$ 为伴随系统 (7.78) 对应的强解. 则对任何满足 $w_\varepsilon + \theta w \in W_{ad}$ (其中 $\theta \in [0,\theta_0]$ 任意, 而 θ_0 为某正数) 的 $w \in L^2(0,T;W)$ 有

$$\int_Q \Big(B^* p_{1\varepsilon} + \frac{\partial f}{\partial w}\big(x,t,w_\varepsilon\big) - 2(w_\varepsilon - \bar{w})\Big) w \,\mathrm{d}x\,\mathrm{d}t \leqslant 0.$$

证明　令 $y_{1\varepsilon w}$ 为 (7.69) 对应于 $w_\varepsilon + \theta w$ 的强解. 易验证在 $\mathcal{C}\big([0,T];X\big)$ 的范数拓扑下当 $\theta \downarrow 0$ 时有 $y_{1\varepsilon w} \to y_{1\varepsilon}$, 且有表示

$$\frac{y_{1\varepsilon w}(x,t) - y_{1\varepsilon}(x,t)}{\theta} = z_{1\varepsilon}(x,t) \tag{7.79}$$

在 $\theta > 0$ 时对 $(x,t) \in Q$ 几乎处处成立, 其中 $z_{1\varepsilon}$ 是下述系统的强解:

$$\begin{cases} \dfrac{\partial z_1}{\partial t} + A z_1 = Bw, & \text{在 } Q \text{ 中几乎处处成立}, \\ z_1(x,0) = 0, & x \in \Omega, \\ z_1(s,t) = 0, & (s,t) \in \Sigma. \end{cases}$$

定义极限

$$\delta_1 := \limsup_{\theta \downarrow 0} \int_Q \frac{g\big(x,t,y_{1\varepsilon w}(x,t) + \bar{y}_2(x,t)\big) - g\big(x,t,y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big)}{\theta}\,\mathrm{d}x\,\mathrm{d}t,$$

$$\delta_2 := \limsup_{\theta \downarrow 0} \int_Q \frac{\varepsilon\alpha_\varepsilon^2\big(y_{1\varepsilon w}(x,t) + \bar{y}_2(x,t)\big) - \varepsilon\alpha_\varepsilon^2\big(y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big)}{\theta}\,\mathrm{d}x\,\mathrm{d}t,$$

并对其中的被积项应用中值定理得

$$\delta_1 = \limsup_{\theta\downarrow 0}\int_Q \frac{\partial g}{\partial y}\Big(x,t,y_{1\varepsilon}+\bar{y}_2+\theta_1(y_{1\varepsilon w}-y_{1\varepsilon})\Big)\frac{y_{1\varepsilon w}(x,t)-y_{1\varepsilon}(x,t)}{\theta}\,\mathrm{d}x\,\mathrm{d}t,$$

$$\delta_2 = \varepsilon\limsup_{\theta\downarrow 0}\int_Q \big(\alpha_\varepsilon(y_{1\varepsilon w}+\bar{y}_2)+\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\big)\alpha_\varepsilon'\big(y_{1\varepsilon}+\bar{y}_2+\theta_2(y_{1\varepsilon w}-y_{1\varepsilon})\big)$$
$$\times\frac{y_{1\varepsilon w}(x,t)-y_{1\varepsilon}(x,t)}{\theta}\,\mathrm{d}x\,\mathrm{d}t,$$

其中 $\theta_1=\theta_1(x,t)$, $\theta_2=\theta_2(x,t)\in[0,1]$ 在 Q 中几乎处处成立. 然后由 (7.79), (H4b), 和 Lebesgue 控制收敛定理得

$$\begin{aligned}&\bigg|\int_Q\Big(\frac{\partial g}{\partial y}\big(x,t,y_{1\varepsilon}+\bar{y}_2+\theta_1(y_{1\varepsilon w}-y_{1\varepsilon})\big)\frac{y_{1\varepsilon w}-y_{1\varepsilon}}{\theta}\\&\quad-\frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2)z_{1\varepsilon}\Big)\,\mathrm{d}x\,\mathrm{d}t\bigg|\\\leqslant&\int_Q\bigg|\frac{\partial g}{\partial y}\big(x,t,y_{1\varepsilon}+\bar{y}_2+\theta_1(y_{1\varepsilon w}-y_{1\varepsilon})\big)\\&\quad-\frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2)\bigg|\cdot|z_{1\varepsilon}|\,\mathrm{d}x\,\mathrm{d}t\to 0\quad(\theta\downarrow 0).\end{aligned}$$

这蕴涵表达式

$$\delta_1=\int_Q\frac{\partial g}{\partial y}\big(x,t,y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)\big)z_{1\varepsilon}(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

进一步注意到 $\alpha_\varepsilon'(\cdot)$ 是连续的且 $|\varepsilon\alpha_\varepsilon'(\cdot)|\leqslant 1$ 和

$$\alpha_\varepsilon(y_{1\varepsilon w}+\bar{y}_2)+\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\in L^2(0,T;X),$$

则由 (7.79) 和上面的计算得

$$\delta_2=2\varepsilon\int_Q\alpha_\varepsilon'\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)\big)\alpha_\varepsilon\big(y_{1\varepsilon}(x,t)+\bar{y}_2(x,t)\big)z_{1\varepsilon}(x,t)\,\mathrm{d}x\,\mathrm{d}t.$$

因为 $w_\varepsilon+\theta w\to w_\varepsilon$ (沿 $L^2(Q)$ 的强拓扑, 其中 $\theta\downarrow 0$) 对任何满足定理条件的 w 成立, 由 (H5b) 和中值定理得

$$\begin{aligned}&\left|\frac{f(x,t,w_\varepsilon+\theta w)-f(x,t,w_\varepsilon)}{\theta}\right|=\left|\frac{\partial f}{\partial w}(x,t,w_\varepsilon+\theta_3\theta w)w\right|\\\leqslant&\,\kappa_1(x,t)|w(x,t)|,\\&\frac{\partial f}{\partial w}(x,t,w_\varepsilon+\theta_3\theta w)w\to\frac{\partial f}{\partial w}(x,t,w_\varepsilon)w\quad\text{在 } Q \text{ 中几乎处处成立},\quad\theta\downarrow 0,\end{aligned}$$

其中 $\theta_3=\theta_3(x,t)\in[0,1]$ 在 Q 中几乎处处成立. 因此 Lebesgue 控制收敛定理蕴涵

$$\int_Q\frac{\partial f}{\partial w}(x,t,w_\varepsilon+\theta_3\theta w)w\,\mathrm{d}x\,\mathrm{d}t\to\int_Q\frac{\partial f}{\partial w}(x,t,w_\varepsilon)w\,\mathrm{d}x\,\mathrm{d}t\quad(\theta\downarrow 0).$$

现在应用 $(w_\varepsilon, y_{1\varepsilon})$ 在问题 $(\mathrm{P}_{1\varepsilon})$ 中的最优性, 有

$$
\begin{aligned}
0 \geqslant & \limsup_{\theta\downarrow 0} \frac{J_{1\varepsilon}(w_\varepsilon+\theta w, y_{1\varepsilon w}) - J_{1\varepsilon}(w_\varepsilon, y_{1\varepsilon})}{\theta} \\
\geqslant & \limsup_{\theta\downarrow 0} \int_Q \left[\frac{g(x,t,y_{1\varepsilon w}+\bar{y}_2) - g(x,t,y_{1\varepsilon}+\bar{y}_2)}{\theta} \right. \\
& \left. + \frac{f(x,t,w_\varepsilon+\theta w) - f(x,t,w_\varepsilon)}{\theta} \right] \mathrm{d}x\,\mathrm{d}t \\
& - \limsup_{\theta\downarrow 0} \int_Q \frac{(w_\varepsilon+\theta w-\bar{w})^2 - (w_\varepsilon-\bar{w})^2}{\theta} \mathrm{d}x\,\mathrm{d}t \\
& -\varepsilon \limsup_{\theta\downarrow 0} \int_Q \frac{\alpha_\varepsilon^2(y_{1\varepsilon w}+\bar{y}_2) - \alpha_\varepsilon^2(y_{1\varepsilon}+\bar{y}_2)}{\theta} \mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

综上所述, 则得不等式

$$
\begin{aligned}
0 \geqslant & \int_Q \left(\frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2) - 2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2) \right) z_{1\varepsilon}\,\mathrm{d}x\,\mathrm{d}t \\
& + \int_Q \left(\frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2(w_\varepsilon-\bar{w}) \right) w\,\mathrm{d}x\,\mathrm{d}t.
\end{aligned}
$$

最后把 (7.70) 的解 $p_{1\varepsilon}$ 替换进这个不等式并分部积分, 就完成定理的证明. △

推论 7.42(最差扰动逐点形式的次最优条件) 对任意 $\varepsilon > 0$, 问题 $(\mathrm{P}_{1\varepsilon})$ 中的极大扰动 w_ε 满足砰–砰关系, 即

$$
w_\varepsilon(x,t) = c \quad \text{和} \quad w_\varepsilon(x,t) = d
$$

分别在集合

$$
\left\{ (x,t)\in Q \mid (B^* p_{1\varepsilon})(x,t) + \frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2\big(w_\varepsilon(x,t) - \bar{w}(x,t)\big) < 0 \right\}
$$

和

$$
\left\{ (x,t)\in Q \mid (B^* p_{1\varepsilon})(x,t) + \frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2\big(w_\varepsilon(x,t) - \bar{w}(x,t)\big) > 0 \right\}
$$

上几乎处处成立, 其中 $p_{1\varepsilon}$ 是伴随系统 (7.79) 的对应解.

证明 对任意 $w \in W_{ad}$ 取 $\widetilde{w} := w - w_\varepsilon$, 则 $w_\varepsilon + \theta\widetilde{w} = (1-\theta)w_\varepsilon + \theta w \in W_{ad}$ 在 $\theta \in [0,1]$ 时成立. 在定理 7.41 的最优条件中以 $\widetilde{w}$ 替换 w, 就有

$$
\int_Q \left(B^* p_{1\varepsilon} + \frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2(w_\varepsilon - \bar{w}) \right) (w - w_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t \leqslant 0
$$

对任意 $w \in W_{ad}$ 成立, 这蕴涵所需的砰–砰关系. △

7.4.4 最差扰动的次最优控制

本小节研究 7.4.1 小节表述的边界最优控制问题 (P_2). 根据分拆过程, 由 (P_2) 的最优解可以找到最差扰动下原极小极大问题 (P) 的最优边界控制.

所论的问题 (P_2) 是抛物系统的边界最优控制问题, 它具有在 Dirichlet 边界条件中起作用的逐点/硬控制约束, 和由分拆过程生成的移动状态约束. 为消除/逼近这种状态约束, 这里建立了一种罚函数技术, 它能提供原极小极大问题有用的次最优信息.

令 $\alpha(\cdot)$ 为上一小节定义的极大单调算子, $\alpha_\varepsilon(\cdot)$ 为 $\alpha(\cdot)$ 具有形式 (7.73) 的光滑逼近. 对任意 $\varepsilon>0$, 考虑如下 (P_2) 的一族逼近问题:

$$
\begin{aligned}
(P_{2\varepsilon})\ \min\ J_{2\varepsilon}(u,y_2) := &\int_Q g\big(x,t,\bar{y}_1(x,t)+y_2(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t\\
&+\int_\Sigma h\big(s,t,u(s,t)\big)\,\mathrm{d}s\,\mathrm{d}t+\big\|u-\bar{u}\big\|^p_{L^p(0,T;U)}\\
&+\varepsilon\big\|\alpha_\varepsilon(\bar{y}_1+y_2)\big\|^2_{L^2(0,T;X)},
\end{aligned}
$$

其中 $u\in U_{ad}$ 并满足系统 (7.70).

注意 (7.70) 的解应理解为适度解, 即满足

$$
y_2(t)=\mathcal{L}u:=A\int_0^t S(t-\tau)Du(\tau)\,\mathrm{d}\tau\quad(\forall t\in[0,T])
$$

的 $y_2\in\mathcal{C}\big([0,T];X\big)$. 下述结果证明了对任意 $\varepsilon>0$ 极小化逼近问题 $(\mathrm{P}_{2\varepsilon})$ 最优解的存在性.

定理 7.43(Dirichlet 问题的逼近的最优解的存在性) 对任意 $\varepsilon>0$, 逼近问题 $(\mathrm{P}_{2\varepsilon})$ 至少有一个最优解对 $(u_\varepsilon,y_{2\varepsilon})\in U_{ad}\times\mathcal{C}\big([0,T];X\big)$.

证明 首先注意到, 每个问题 $(\mathrm{P}_{2\varepsilon})$ 都具有由原极小极大问题 (P) 的给定最优解 $(\bar{u},\bar{w})$ 生成的可行解 $(\bar{u},\bar{y}_2)$. 令 (u,y_2) 为 $(\mathrm{P}_{2\varepsilon})$ 的任意可行解对. 由假设 (H4a) 和 (H6a) 得积分和

$$
\int_Q g\big(x,t,\bar{y}_1(x,t)+y_2(x,t)\big)\,\mathrm{d}x\,\mathrm{d}t+\int_\Sigma h\big(s,t,u(s,t)\big)\,\mathrm{d}s\,\mathrm{d}t
$$

对于 $u\in U_{ad}$ 和 Dirichlet 系统 (7.70) 对应的轨道 y_2 一致下有界. 为估计分布系统 (7.69) 的给定轨道 $\bar{y}_1$, 应用定义 7.31 里适度解的表示中的指数半群估计 (7.65), 则得

$$
\big\|\bar{y}_1(t)\big\|_X\leqslant M_1\Big(\mathrm{e}^{-\omega t}\|y_0\|_X+\frac{\big(\|B\|\max\{|c|,d\}+\|\vartheta\|_\infty\big)\sqrt{\operatorname{mes}(\Omega)}}{\omega}\big(1-\mathrm{e}^{-\omega t}\big)\Big).
$$

然后以类似定理 7.38 证明的推理, 即可得罚函数项 $\|\alpha_\varepsilon(\bar{y}_1+y_2)\|_{L^2(0,T;X)}$ 的一致有界性. 这给出费用函数 $J_{2\varepsilon}(u,y_2)$ 的一致下有界性, 进而有对任意 $\varepsilon>0$ 问题 $(\mathrm{P}_{2\varepsilon})$ 中的下确界 $\inf J_{2\varepsilon}(u,y_2)$ 的有限性.

进一步, 考虑到系统 (7.70) 对应于任何给定容许控制 $u\in U_{ad}$ 的适度解 y_2 的存在性和唯一性, 则可把 $(\mathrm{P}_{2\varepsilon})$ 中的费用函数 $J_{2\varepsilon}$ 作为 $u\in U_{ad}\subset L^p(0,T;U)$ 的函数, 其中最后这个空间具有弱拓扑. 现在由推论 7.34 和定理 7.35 中适度解的正则性/收敛性, 以及被积项 h 对于 u 的凸性, 得出 $(\mathrm{P}_{2\varepsilon})$ 中的费用函数在 $L^p(0,T;U)$ 中在弱紧集 U_{ad} 上是弱下半连续的; 比照定理 7.38 的证明. 这样在所论的拓扑下 $(\mathrm{P}_{2\varepsilon})$ 最优解的存在性就可由经典的 Weierstrass 定理导出. △

下面建立逼近问题 $(\mathrm{P}_{2\varepsilon})$ 的最优解对给定的状态约束问题 (P_2) 的最优解 $(\bar{u},\bar{y}_2)$ 的强收敛性.

定理 7.44(Dirichlet 边界控制逼近问题的强收敛性) 设 $(\bar{u},\bar{y}_2)$ 为状态约束问题 (P_2) 的给定最优解, $\{(u_\varepsilon,y_{2\varepsilon})\}$ 为逼近问题 $(\mathrm{P}_{2\varepsilon})$ 的最优解序列. 则存在正常数 ε 组成的序列满足

$$\begin{aligned}&u_\varepsilon\to\bar{u}\quad\big(L^p(0,T;U)\text{ 中的强收敛}\big),\\&y_{2\varepsilon}\to\bar{y}_2\quad\big(\mathcal{C}\big([0,T];X\big)\text{ 中的强收敛}\big),\\&J_{2\varepsilon}(u_\varepsilon,y_{2\varepsilon})\to J_2(\bar{u},\bar{y}_2)\quad(\varepsilon\downarrow 0).\end{aligned}$$

证明 由 $(u_\varepsilon,y_{2\varepsilon})$ 对 $(\mathrm{P}_{2\varepsilon})$ 的最优性和 $(\bar{u},\bar{y}_2)$ 的可行性, 得

$$J_{2\varepsilon}(u_\varepsilon,y_{2\varepsilon})\leqslant J_{2\varepsilon}(\bar{u},\bar{y}_2)=J_2(\bar{u},\bar{y}_2),\quad\forall\varepsilon>0.\tag{7.80}$$

这特别地蕴涵, 存在不依赖于 ε 的 $M>0$, 满足

$$\varepsilon\big\|\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\big\|^2_{L^2(0,T;X)}\leqslant M,\quad\varepsilon\big\|\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\big\|_{L^2(0,T;X)}\to 0\quad(\varepsilon\downarrow 0).$$

根据 U_{ad} 在 $L^p(0,T;U)$ 中的弱紧性, 找到函数 $\widetilde{u}\in U_{ad}$ 和 $\{u_\varepsilon\}$ 的子序列使得

$$u_\varepsilon\to\widetilde{u}\quad(\varepsilon\downarrow 0,\quad L^p(0,T;U)\text{ 中的弱收敛}).$$

以 $\widetilde{y}_2$ 记对应于 $\widetilde{u}$ 的 (7.70) 的 (唯一) 适度解. 由定理 7.35, 选取 $\varepsilon\downarrow 0$ 的子序列满足

$$y_{2\varepsilon}(x,t)\to\widetilde{y}_2(x,t)\quad(\text{在 }Q\text{ 中几乎处处}),$$

其中 p 足够大. 然后沿着定理 7.40 的证明思路并利用引理 7.39, 可以验证状态约束

$$a\leqslant\bar{y}_1(x,t)+\widetilde{y}_2(x,t)\leqslant b\quad(\text{在 }Q\text{ 中几乎处处})$$

的正确性. 这保证了 $(\widetilde{u}, \widetilde{y}_2)$ 是状态约束问题 (P_2) 的一个可行解, 从而

$$J_2(\widetilde{u}, \widetilde{y}_2) \geqslant J_2(\bar{u}, \bar{y}_2).$$

考虑到费用函数 J_2 在 $L^p(0,T;U)$ 中的弱拓扑下在 U_{ad} 上的下半连续性, 当 $\varepsilon \downarrow 0$ 时在 (7.80) 中取极限 (比照定理 7.36 的证明), 则得

$$\lim_{\varepsilon\downarrow 0} \left\|u_\varepsilon - \bar{u}\right\|^p_{L^p(0,T;U)} = 0, \quad \lim_{\varepsilon\downarrow 0} \varepsilon\left\|\alpha_\varepsilon(\bar{y}_1 + y_{2\varepsilon})\right\|^2_{L^2(0,T;X)} = 0. \tag{7.81}$$

(7.81) 中的第一个等式意味着 $L^p(0,T;U)$ 中的强收敛 $u_\varepsilon \to \bar{u}$ $(\varepsilon \downarrow 0)$. 由定理 7.33, 在 p 足够大时, 这蕴涵 $\mathcal{C}([0,T];X)$ 中的强收敛 $y_{2\varepsilon} \to \bar{y}_2$. 最后, 根据 $J_{2\varepsilon}$ 的罚函数构造, 定理中的值收敛性可由 (7.81) 的第二个等式得出. △

接下来在下述假设下推导逼近问题必要最优条件. 该假设和 7.4.3 小节中的是平行的条件 (请比照那里的讨论).

(H6b) $h(s,t,u)$ 对于 u 连续可微且导数对于 (s,t) 可测. 进一步, 存在非负函数 $\gamma_1 \in L^q(\Sigma)$ (其中 $1/p + 1/q = 1$), 满足估计

$$\left|\frac{\partial h}{\partial u}(s,t,u)\right| \leqslant \gamma_1(s,t), \quad \text{a.e. } (s,t) \in \Sigma \quad (\text{当} u \in [\mu, \nu]).$$

设 $(u_\varepsilon, y_{2\varepsilon})$ 为逼近问题 $(\mathrm{P}_{2\varepsilon})$ 的一个最优解, 其中 $\varepsilon > 0$ 为任意固定常数. 考虑控制 u_ε 的具有形式 $u_\varepsilon + \theta u \in U_{ad}$ 的可行变分, 其中 $u \in L^p(0,T;U)$, $\theta \in [0, \theta_0]$, 而 $\theta_0 > 0$ 为某常数. 以 $y_{2\varepsilon u}$ 记对应于 $u_\varepsilon + \theta u$ 的 (7.70) 的适度解并考虑定义为

$$\varphi(\theta) := J_{2\varepsilon}(u_\varepsilon + \theta u, y_{2\varepsilon u})$$

的函数 $\varphi\colon [0,\theta_0] \to \mathbb{R}$. 该函数显然在 $\theta = 0$ 时达到极小. 另外从适度解的定义可知

$$\begin{aligned} &y_{2\varepsilon u} \to y_{2\varepsilon} \quad (\theta \downarrow 0,\ \mathcal{C}([0,T]; H^{1/2-\varepsilon}(\Omega)) \text{ 中的强收敛}), \\ &\frac{y_{2\varepsilon u}(x,t) - y_{2\varepsilon}(x,t)}{\theta} = \mathcal{L}u, \quad \text{a.e. } (x,t) \in Q \quad (\text{当}\theta > 0), \end{aligned}$$

其中 p 足够大.

现在可以导出 Dirichlet 边界控制逼近问题 $(\mathrm{P}_{2\varepsilon})$ 的必要最优条件了. 首先得到积分形式, 然后作为推论得到其 (逐点) 砰–砰原理形式. 如前述, 由定理 7.44 的强收敛性, 这些结果提供了状态约束问题 (P_2) 的次最优条件.

定理 7.45(最差扰动下 Dirichlet 边界控制的次最优条件) 对固定的 $\varepsilon > 0$, 设 $(u_\varepsilon, y_{2\varepsilon})$ 为逼近问题 $(\mathrm{P}_{2\varepsilon})$ 的一个最优解, 并令

$$\mathcal{L}^*\colon \mathcal{C}([0,T];X)^* \to L^q(0,T;U)$$

为 (7.72) 中定义的适度解映射 $\mathcal{L}$ 的伴随算子. 则

$$0\leqslant\int_{\Sigma}\Big[\mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar y_1+y_{2\varepsilon})+2\varepsilon\alpha_\varepsilon'(\bar y_1+y_{2\varepsilon})\alpha_\varepsilon(\bar y_1+y_{2\varepsilon})\Big)$$
$$+\frac{\partial h}{\partial u}(s,t,u_\varepsilon)\Big]u\,\mathrm{d}s\,\mathrm{d}t+2p\int_0^T\big\|u_\varepsilon-\bar u\big\|_U^{p-2}\Big(\int_\Gamma(u_\varepsilon-\bar u)u\,ds\Big)\,\mathrm{d}t$$

对任意满足

$$u_\varepsilon+\theta u\in U_{ad}\quad(\forall\theta\in[0,\theta_0])$$

的 $u\in L^p(0,T;U)$ 成立, 其中 $\theta_0>0$ 为某常数.

证明　由于上述函数 φ 显然在 $\theta=0$ 处达到最小, 应用经典的中值定理可得关系:

$$0\leqslant\liminf_{\theta\downarrow0}\frac{\varphi(\theta)-\varphi(0)}{\theta}+\liminf_{\theta\downarrow0}\frac{1}{\theta}\Big[\int_Q\Big(g(x,t,\bar y_1+y_{2\varepsilon u})-g(x,t,\bar y_1+y_{2\varepsilon})\Big)\,\mathrm{d}x\mathrm{d}t$$
$$+\int_\Sigma\Big(h(s,t,u_\varepsilon+\theta u)-h(s,t,u_\varepsilon)\Big)\,ds\,dt+\Big(\big\|u_\varepsilon+\theta u-\bar u\big\|^p_{L^p(0,T;U)}$$
$$-\big\|u_\varepsilon-\bar u\big\|^p_{L^p(0,T;U)}\Big)+\varepsilon\Big(\big\|\alpha_\varepsilon(\bar y_1+y_{2\varepsilon u})\big\|^2_{L^2(0,T;X)}-\big\|\alpha_\varepsilon(\bar y_1+y_{2\varepsilon})\big\|^2_{L^2(0,T;X)}\Big)\Big]$$
$$=\liminf_{\theta\downarrow0}\frac{1}{\theta}\Big[\int_Q\frac{\partial g}{\partial y}\Big(x,t,\bar y_1+y_{2\varepsilon}+\theta_1(y_{2\varepsilon u}-y_{2\varepsilon})\Big)(y_{2\varepsilon u}-y_{2\varepsilon})\,\mathrm{d}x\,\mathrm{d}t$$
$$+\int_\Sigma\frac{\partial h}{\partial u}(s,t,u_\varepsilon+\theta_2\theta u)\theta u\,\mathrm{d}s\,\mathrm{d}t$$
$$+\int_0^T\Big(\big\|u_\varepsilon+\theta u-\bar u\big\|_U^{p-2}+\cdots+\big\|u_\varepsilon-\bar u\big\|_U^{p-2}\Big)\Big(\int_\Gamma\theta u(2u_\varepsilon-2\bar u+\theta u)\mathrm{d}s\Big)\mathrm{d}t$$
$$+\varepsilon\int_Q\Big(\alpha_\varepsilon(\bar y_1+y_{2\varepsilon u})+\alpha_\varepsilon(\bar y_1+y_{2\varepsilon})\Big)\alpha_\varepsilon'\Big(\bar y_1+y_{2\varepsilon}$$
$$+\theta_3(y_{2\varepsilon u}-y_{2\varepsilon})\Big)(y_{2\varepsilon u}-y_{2\varepsilon})\,\mathrm{d}x\,\mathrm{d}t\Big],$$

其中 $\theta_i=\theta_i(x,t)\in[0,1]$ (在 Q 中几乎处处成立, $i=1,2,3$). 注意到 $\theta_i(y_{2\varepsilon u}-y_{2\varepsilon})\to 0$ ($\theta\downarrow0$, $L^2(Q)$ 中的强收敛, $i=1,2,3$) 和

$$\alpha_\varepsilon(\bar y_1+y_{2\varepsilon u})+\alpha_\varepsilon(\bar y_1+y_{2\varepsilon})\in L^2(0,T;X),$$

然后与 7.4.3 小节类似, 由 (H4b) 和 (H6b) 以及 Lebesgue 控制收敛定理, 得不等式

$$0\leqslant\int_Q\Big(\frac{\partial g}{\partial y}(x,t,\bar y_1+y_{2\varepsilon})+2\varepsilon\alpha_\varepsilon'(\bar y_1+y_{2\varepsilon})\alpha_\varepsilon(\bar y_1+y_{2\varepsilon})\Big)\mathcal{L}u\,\mathrm{d}x\,\mathrm{d}t$$
$$+\int_\Sigma\frac{\partial h}{\partial u}(s,t,u_\varepsilon)u\,\mathrm{d}s\,\mathrm{d}t+2p\int_0^T\big\|u_\varepsilon-\bar u\big\|_U^{p-2}\Big(\int_\Gamma(u_\varepsilon-\bar u)u\,\mathrm{d}s\Big)\,\mathrm{d}t.$$

这蕴涵定理中给出的必要最优条件. △

由定理 7.45 的必要最优条件易得扰动问题 $(P_{2\varepsilon})$ 的砰–砰形式的最优控制. 注意下面给出的砰–砰关系对边界 Σ 上的奇点 (s,t) 的集合上的控制没有提供任何信息, 这里的奇点是指在该点上定理 7.45 中积分不等式的对应被积项为零.

推论 7.46(Dirichlet 边界控制的砰–砰次最优条件) 对任意 $\varepsilon > 0$, $(P_{2\varepsilon})$ 的最优控制 u_ε 满足砰–砰关系, 即

$$u_\varepsilon(s,t) = \mu \quad \text{和} \quad u_\varepsilon(s,t) = \nu$$

分别在集合

$$\Big\{(s,t) \in \Sigma \mid \mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y}_1 + y_{2\varepsilon}) + 2\varepsilon\alpha'_\varepsilon(\bar{y}_1 + y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1 + y_{2\varepsilon})\Big) \\ + \frac{\partial h}{\partial u}(s,t,u_\varepsilon) + 2p\big\|u_\varepsilon - \bar{u}\big\|_U^{p-2}(u_\varepsilon - \bar{u}) < 0\Big\}$$

和

$$\Big\{(s,t) \in \Sigma \mid \mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y}_1 + y_{2\varepsilon}) + 2\varepsilon\alpha'_\varepsilon(\bar{y}_1 + y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1 + y_{2\varepsilon})\Big) \\ + \frac{\partial h}{\partial u}(s,t,u_\varepsilon) + 2p\big\|u_\varepsilon - \bar{u}\big\|_U^{p-2}(u_\varepsilon - \bar{u}) > 0\Big\}$$

上几乎处处成立.

证明 在定理 7.45 中对任意 $v \in U_{ad}$ 置 $u := v - \bar{u}$. 取 $\theta_0 = 1$, 则 $u_\varepsilon + \theta u = (1-\theta)u_\varepsilon + \theta v \in U_{ad}$ 对任意 $\theta \in [0,1]$ 成立. 因此所给的砰–砰条件可直接由定理中的积分不等式条件得出. △

7.4.5 状态约束下的必要最优条件

本节也是本章的最后建立所需的极限过程以导出原极小极大控制问题 (P) 必要最优条件. 利用分拆的过程和前面小节证明的强收敛结果, 这种极限过程基于对逼近问题 $(P_{1\varepsilon})$ 和 $(P_{2\varepsilon})$ 的必要最优条件取极限. 首先总结一下对原问题 (P) 的给定最优解 $(\bar{u},\bar{w})$ 建立的逼近和次最优结果.

定理 7.47(极小极大解的次最优条件) 设 $(\bar{u},\bar{w})$ 为极小极大控制问题 (P) 的最优解, $\bar{y}$ 为对应的系统 (7.61) 在假设 (H1)~(H6) 下的适度轨道 (其中 $p > 0$ 足够大). 则对任意 $\varepsilon > 0$, 分别存在问题 $(P_{1\varepsilon})$ 和 $(P_{2\varepsilon})$ 的最优解 $(w_\varepsilon, y_{1\varepsilon})$ 和 $(u_\varepsilon, y_{2\varepsilon})$, 它们强逼近 $(\bar{u},\bar{w})$, 即在 $L^p(0,T;U) \times L^2(0,T;W) \times \mathcal{C}\big([0,T];X\big)$ 中

$$(u_\varepsilon, w_\varepsilon, y_{1\varepsilon} + y_{2\varepsilon}) \to (\bar{u},\bar{w},\bar{y}) \quad (\varepsilon \downarrow 0),$$

并且满足定理 7.41 和定理 7.45 中对应的必要最优条件.

分析一下定理 7.41 和定理 7.45 的必要条件, 可以看出在 $\varepsilon \downarrow 0$ 时取极限需要逼近项 $\varepsilon\alpha_\varepsilon'(\cdot)\alpha_\varepsilon(\cdot)$ 的一致有界性. 没有额外假设这是不成立的. 现在对原极小极大问题中的状态约束加上看起来是最合适的“规范条件”以建立极限过程. 下述中 $\|\cdot\|_\infty$ 表示 $L^\infty(Q)$ 的范数.

(**CQ1**) 存在 $\widetilde{w} \in W_{ad}$ 和 $\eta_1 > 0$ 使得, 对任意满足 $\|\zeta\|_\infty \leqslant 1$ 的 $\zeta \in L^\infty(Q)$ 有

$$a \leqslant \widetilde{y}_1(x,t) + \bar{y}_2(x,t) + \eta_1\zeta(x,t) \leqslant b \quad (\text{在 } Q \text{ 中几乎处处成立}),$$

其中 $\widetilde{y}_1$ 是系统 (7.69) 由 $\widetilde{w}$ 生成的 (唯一) 强解.

(**CQ2**) 存在 $\widetilde{u} \in U_{ad}$ 和 $\eta_2 > 0$ 使得, 对任意满足 $\|\zeta\|_\infty \leqslant 1$ 的 $\zeta \in L^\infty(Q)$ 有

$$a \leqslant \bar{y}_1(x,t) + \widetilde{y}_2(x,t) + \eta_2\zeta(x,t) \leqslant b \quad (\text{在 } Q \text{ 中几乎处处成立}),$$

其中 $\widetilde{y}_2$ 为系统 (7.70) 由 $\widetilde{y}$ 生成的 (唯一) 适度解.

值得注意的是, 上面的规范条件和在所论可行解空间中凸规划的经典 Slater 约束规范在无穷维空间中的版本是不同的. 特别地, 它们不蕴涵问题 (P_1) 和 (P_2) 可行解轨道 y_1 和 y_2 组成的集合分别在 $W^{1,2}\big([0,T];X\big)$ 和 $\mathcal{C}\big([0,T];X\big)$ 中具有非空内部.

下面的引理提供了所需的逼近项的一致估计, 这对极限过程的建立至关重要.

引理 7.48(约束规范下的一致估计)　设 $(\bar{u},\bar{w},\bar{y})$, $(w_\varepsilon, y_{1\varepsilon})$, 和 $(u_\varepsilon, y_{2\varepsilon})$ 满足定理 7.47 的条件. 再假设约束规范 (CQ1) 和 (CQ2) 成立. 则存在与 ε 无关的常数 $C>0$ 使得对任意 $\varepsilon>0$ 有估计

$$\big\|\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\big\|_1 \leqslant C, \quad \big\|\varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\big\|_1 \leqslant C,$$

其中 $\|\cdot\|_1$ 表示 $L^1(Q)$ 的范数.

证明　设 $\widetilde{w}$ 满足约束规范条件 (CQ1). 考虑扰动

$$w := \widetilde{w} - w_\varepsilon, \quad \forall \varepsilon > 0$$

并将其替换进定理 7.41 证明中的最后一个不等式. 根据 $\alpha_\varepsilon(\cdot)$ 单调性, 有估计

$$\begin{aligned}0 \geqslant & \int_Q \frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2)(\widetilde{y}_1 - y_{1\varepsilon})\,\mathrm{d}x\,\mathrm{d}t \\ & + \int_Q \Big(\frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2(w_\varepsilon - \bar{w})\Big)(\widetilde{w} - w_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t \\ & -2\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)(\widetilde{y}_1 - y_{1\varepsilon})\,\mathrm{d}x\,\mathrm{d}t\end{aligned}$$

$$
\begin{aligned}
&= \int_Q \frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2)(\widetilde{y}_1 - y_{1\varepsilon})\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_Q \Big(\frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2(w_\varepsilon - \bar{w})\Big)(\widetilde{w} - w_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\big(\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2) - \alpha_\varepsilon(\widetilde{y}_1+\bar{y}_2+\eta_1\zeta)\big) \\
&\quad \times (y_{1\varepsilon}+\bar{y}_2-\widetilde{y}_1-\bar{y}_2-\eta_1\zeta)\,\mathrm{d}x\,\mathrm{d}t \\
&\quad +2\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\eta_1\zeta\,\mathrm{d}x\,\mathrm{d}t \\
&\geqslant \int_Q \frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar{y}_2)(\widetilde{y}_1 - y_{1\varepsilon})\,\mathrm{d}x\,\mathrm{d}t \\
&\quad + \int_Q \Big(\frac{\partial f}{\partial w}(x,t,w_\varepsilon) - 2(w_\varepsilon - \bar{w})\Big)(\widetilde{w} - w_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t \\
&\quad +2\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\eta_1\zeta\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
$$

对任何满足 $|\zeta\|_\infty \leqslant 1$ 的 $\zeta \in L^\infty(Q)$ 成立. 由假设 (H4b), (H5b) 和定理 7.40 易知, 存在与 ε 无关的常数 $C > 0$ 保证所需的一致有界性

$$
\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\zeta\,\mathrm{d}x\,\mathrm{d}t \leqslant C
$$

对任意 $\varepsilon > 0$ 和满足 $\|\zeta\|_\infty \leqslant 1$ 的 $\zeta \in L^\infty(Q)$ 成立. 这就显然蕴涵引理中的第一个估计.

对第二个估计, 取满足约束规范条件 (CQ2) 的 $\widetilde{u}$ 并把控制

$$
u := \widetilde{u} - u_\varepsilon, \quad \forall \varepsilon > 0
$$

替换进定理 7.45 证明中的最后一个不等式. 再一次由 $\alpha_\varepsilon(\cdot)$ 的单调性可得

$$
\begin{aligned}
0 &\leqslant \int_Q \frac{\partial g}{\partial y}(x,t,\bar{y}_1+y_{2\varepsilon})\mathcal{L}(\widetilde{u}-u_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma \frac{\partial h}{\partial u}(s,t,u_\varepsilon)(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad +2p\int_0^T \|u_\varepsilon(t)-\bar{u}(t)\|_U^{p-2}\Big(\int_\Gamma (u_\varepsilon-\bar{u})(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\Big)\,\mathrm{d}t \\
&\quad +2\int_Q \varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})(\mathcal{L}\widetilde{u}-\mathcal{L}u_\varepsilon)\,\mathrm{d}x\,\mathrm{d}t \\
&\leqslant \int_Q \frac{\partial g}{\partial y}(x,t,\bar{y}_1+y_{2\varepsilon})(\widetilde{y}_2-y_{2\varepsilon})\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma \frac{\partial h}{\partial u}(s,t,u_\varepsilon)(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\,\mathrm{d}t \\
&\quad +2p\int_0^T \|u_\varepsilon(t)-\bar{u}(t)\|_U^{p-2}\Big(\int_\Gamma (u_\varepsilon-\bar{u})(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\Big)\,\mathrm{d}t
\end{aligned}
$$

$$
\begin{aligned}
&-2\int_Q \varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\big(\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})-\alpha_\varepsilon(\bar{y}_1+\widetilde{y}_2+\eta_2\zeta)\big)\\
&\times(\bar{y}_1+y_{2\varepsilon}-\bar{y}_1-\widetilde{y}_2-\eta_2\zeta)\,\mathrm{d}x\,\mathrm{d}t\\
&-2\eta_2\int_Q \varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\zeta\,\mathrm{d}x\,\mathrm{d}t\\
\leqslant&\int_Q \frac{\partial g}{\partial y}(x,t,\bar{y}_1+y_{2\varepsilon})(\widetilde{y}_2-y_{2\varepsilon})\,\mathrm{d}x\,\mathrm{d}t+\int_\Sigma \frac{\partial h}{\partial u}(s,t,u_\varepsilon)(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\,\mathrm{d}t\\
&+2p\int_0^T \|u_\varepsilon(t)-\bar{u}(t)\|_U^{p-2}\Big(\int_\Gamma(u_\varepsilon-\bar{u})(\widetilde{u}-u_\varepsilon)\,\mathrm{d}s\Big)\,\mathrm{d}t\\
&-2\eta_2\int_Q \varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\zeta\,\mathrm{d}x\,\mathrm{d}t
\end{aligned}
$$

对任意满足 $\|\zeta\|_\infty\leqslant 1$ 的 $\zeta\in L^\infty(Q)$ 成立. 根据假设 (H4b), (H6b) 和定理 7.45 知, 存在不依赖于 $\varepsilon>0$ 的 $C>0$ 使得

$$
\int_Q \varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\zeta\,\mathrm{d}x\,\mathrm{d}t\leqslant C
$$

对任意 $\varepsilon>0$ 和满足 $\|\zeta\|_\infty\leqslant 1$ 的 $\zeta\in L^\infty(Q)$ 成立. 这就蕴涵引理中的第二个估计, 从而完成证明. △

给定极小极大问题 (P) 的最优轨道, 定义集合

$$
Q_{ab}:=\big\{(x,t)\in Q\mid \bar{y}(x,t)=a\ \text{或}\ \bar{y}(x,t)=b\big\},
$$

其中约束 (7.62) 有效. 该集合对引理 7.48 中估计的函数 (看成对偶空间 $L^\infty(Q)^*$ 中的元) 的拓扑极限刻画是非常重要的. 众所周知, 空间 $L^\infty(Q)^*$ 可以等同于定义于 Q 的子集上的有界可加函数组成的空间 $ba(Q)$. 这些可加函数有时称为广义测度, 并且在 Lebesgue 测度为零的集合上取零值. 这意味着对任意 $\varLambda\in L^\infty(Q)^*$ 存在一个唯一测度 $\lambda\in ba(Q)$ 满足

$$
\varLambda(\beta)=\int_Q \beta\,\lambda(\mathrm{d}x\,\mathrm{d}t),\quad \forall\beta\in L^\infty(Q).
$$

下面的讨论不区别空间 $L^\infty(Q)^*$ 和 $ba(Q)$, 即不区分上述关系中的 $\varLambda$ 和 λ. 对任意 $\lambda\in L^\infty(Q)^*$, 考虑其支集 $\operatorname{supp}\lambda$, 即 λ 不为零的集合. 注意和通常一样, 不区分差别为 Q 上 Lebesgue 零测集的支集. $L^\infty(Q)^*$ 上的收敛性总是在拓扑的意义下, 即网收敛. 由于该空间的很强的非序列性质, 这种收敛和序列弱 * 收敛是有本质区别的. 考虑 $L^\infty(Q)^*$ 中的族 $\{\varLambda_\varepsilon\}_{\varepsilon>0}$, 并取其弱 * 收敛的子网, 方便起见, 称为 $\{\varepsilon\}$ 的收敛子网.

引理 7.49(惩罚项的网收敛) 设引理 7.48 的所有假设成立. 则存在满足

$\operatorname{supp}\lambda_i \subset Q_{ab}$ 的 $\lambda_i \in L^\infty(Q)^*$ $(i=1,2)$ 和 $\varepsilon \downarrow 0$ 的子网, 使得沿此子网有

$$2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2) \to \lambda_1 \quad (L^\infty(Q)^* \text{ 中的弱 * 收敛}),$$
$$2\varepsilon\alpha_\varepsilon'(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon}) \to \lambda_2 \quad (L^\infty(Q)^* \text{ 中的弱 * 收敛}).$$

证明 只证第一个收敛关系, 第二个的证明是类似的. 对任意 $\varepsilon>0$, 在 $L^\infty(Q)$ 上定义线性泛函

$$\Lambda_{1\varepsilon}(\beta) := 2\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\beta\,\mathrm{d}x\,\mathrm{d}t, \quad \beta \in L^\infty(Q).$$

由引理 7.48 得

$$\left|\Lambda_{1\varepsilon}(\beta)\right| \leqslant C\left\|\beta\right\|_\infty, \quad \forall\beta \in L^\infty(Q),$$

这保证了每个 $\Lambda_{1\varepsilon}$ 在 $L^\infty(Q)$ 上的连续性, 从而有 $\left\{\Lambda_{1\varepsilon} \mid \varepsilon>0\right\}$ 在空间 $L^\infty(Q)^*$ 中的一致有界性. 由经典的结果, Banach 空间的对偶球是 (拓扑) 弱 * 紧的, 所以可以找到 $\Lambda_1 \in L^\infty(Q)^*$ 和 $\{\varepsilon\}$ 的子网, 沿此子网

$$\lim_{\varepsilon\downarrow 0}\Lambda_{1\varepsilon}(\beta) = \lim_{\varepsilon\downarrow 0} 2\int_Q \varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar{y}_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar{y}_2)\beta\,\mathrm{d}x\,\mathrm{d}t = \Lambda_1(\beta)$$

在 $\beta \in L^\infty(Q)$ 时成立. 这事实上证明了极限测度 $\lambda_1 \in ba(Q)$ (等同于 Λ_1) 的所需收敛性.

余下要证明 $\operatorname{supp}\lambda_1 \subset Q_{ab}$. 注意到, 由状态约束 (7.62) 知集合

$$\left\{(x,t)\in Q \mid \bar{y}(x,t)<a \text{ 或 } \bar{y}(x,t)>b\right\}$$

为零测集. 用反证法假设 λ_1 的支集不含于 Q_{ab}, 则可以找到集合 $\widetilde{Q}$ 满足

$$\begin{cases} \operatorname{mes}\widetilde{Q}>0, \quad \lambda_1(\widetilde{Q})\neq 0, \\ \widetilde{Q}\subset\left\{(x,t)\in Q \mid a<\bar{y}_1(x,t)+\bar{y}_2(x,t)<b\right\}. \end{cases}$$

这显然蕴涵 $\widetilde{Q}\subset\bigcup_{r>0}Q_r$, 其中

$$Q_r := \left\{(x,t)\in Q \mid a+r\leqslant\bar{y}_1(x,t)+\bar{y}_2(x,t)\leqslant b-r\right\}.$$

注意若 $r_1>r_2$, 则 $Q_{r_1}\subset Q_{r_2}$, 因此 $\operatorname{mes}(\widetilde{Q}\cap Q_r)\neq 0$ 对任何小的 $r>0$ 成立. 进一步, 给定 $\delta>0$, 存在 $\tilde{r}>0$ 满足

$$\operatorname{mes}(\widetilde{Q}\backslash Q_{\tilde{r}}) \leqslant \operatorname{mes}\Big(\bigcup_{r>0}Q_r\backslash Q_{\tilde{r}}\Big) < \delta.$$

应用定理 7.40 的强收敛 $y_{1\varepsilon} \to \bar{y}$, 然后由经典的 Egorov 定理, 找到满足 $\mathrm{mes}\big((Q_{\tilde{r}} \cap \widetilde{Q})\backslash Q_\rho\big) < \rho$ 的 $Q_\rho \subset Q_{\tilde{r}} \cap \widetilde{Q}$ 和 $\{y_{1\varepsilon}(x,t)\}$ 的子序列在 Q_ρ 中一致收敛到 $\bar{y}_1(x,t)$. 当 $\rho > 0$ 足够小, 则 $\mathrm{mes}\,(Q_\rho) \neq 0$ 且

$$a < y_{1\varepsilon}(x,t) + \bar{y}_2(x,t) < b$$

对所有小的 ε 在 Q_ρ 中成立. 由 (7.73) 中 $\alpha_\varepsilon(\cdot)$ 的结构, 这蕴涵

$$\varepsilon\alpha'_\varepsilon\big(y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big)\alpha_\varepsilon\big(y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big) = 0$$

对足够小的 $\varepsilon > 0$ 在 Q_ρ 中成立. 还可注意到

$$\widetilde{Q} = (\widetilde{Q} \cap Q_{\tilde{r}}) \cup (\widetilde{Q}\backslash Q_{\tilde{r}}) = Q_\rho \cup \big((\widetilde{Q} \cap Q_{\tilde{r}})\backslash Q_\rho\big) \cup (\widetilde{Q}\backslash Q_{\tilde{r}}).$$

现在考虑满足 $\mathrm{supp}\,\beta \subset \widetilde{Q}$ 的任意 $\beta \in L^\infty(Q)$ 并记

$$\gamma_\varepsilon(x,t) := 2\varepsilon\alpha'_\varepsilon\big(y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big)\alpha_\varepsilon\big(y_{1\varepsilon}(x,t) + \bar{y}_2(x,t)\big)\beta(x,t),$$

则得表示

$$\begin{aligned}\Lambda_{1\varepsilon}(\beta) = &\int_{Q_\rho} \gamma_\varepsilon(x,t)\,\mathrm{d}x\,\mathrm{d}t + \int_{(\widetilde{Q}\cap Q_{\tilde{r}})\backslash Q_\rho} \gamma_\varepsilon(x,t)\,\mathrm{d}x\,\mathrm{d}t \\ &+ \int_{\widetilde{Q}\backslash Q_{\tilde{r}}} \gamma_\varepsilon(x,t)\,\mathrm{d}x\,\mathrm{d}t.\end{aligned}$$

因为 $\gamma_\varepsilon \in L^1(Q)$ 且 δ 足够小, 这蕴涵

$$\Big|\int_{\widetilde{Q}\backslash Q_{\tilde{r}}} \gamma_\varepsilon(x,t)\,\mathrm{d}x\,\mathrm{d}t\Big| < \varepsilon, \quad \forall\varepsilon > 0.$$

由上述关系及引理 7.48 的第一个一致估计, 可找到 $c(\rho) \downarrow 0$ (当 $\rho \to 0$) 使得 $|\Lambda_1(\beta)| \leqslant c(\rho)$ 对任何满足 $\mathrm{supp}\beta \subset \widetilde{Q}$ 的 $\beta \in L^\infty(Q)$ 和足够小的 ρ 成立. 从而 $\Lambda_1(\beta) = 0$ 对该 β 成立, 这与假设矛盾. 证毕. △

现在可以证明具有状态约束的原问题 (P) 的必要最优条件了. 首先建立刻画 (P) 中最差扰动的结果. 给定 $\bar{y} \in \mathcal{C}\big([0,T];X\big)$ 和 $\lambda_1 \in L^\infty(Q)^*$, 考虑伴随系统

$$\begin{cases} \dfrac{\partial p_1}{\partial t} - A^*p_1 = -\dfrac{\partial g}{\partial y}(x,t,\bar{y}) + \lambda_1 \quad (\text{在 } Q \text{ 中几乎处处成立}), \\ p_1(T,x) = 0, \quad x \in \mathrm{cl}\,\Omega, \\ p_1(s,t) = 0, \quad (s,t) \in \Sigma \end{cases} \tag{7.82}$$

并在下述意义下定义其解 $p_1(x,t)$:

$$\int_Q p_1(x,t)\Big(\frac{\partial v}{\partial t} + Av\Big)\,\mathrm{d}x\,\mathrm{d}t = \int_Q \frac{\partial g}{\partial y}\big(x,t,\bar{y}(x,t)\big)v\,\mathrm{d}t\,\mathrm{d}x - \int_Q v\lambda_1(\mathrm{d}x\,\mathrm{d}t),$$
$$\forall v \in W_0^{2,1,\infty}(Q).$$

下面的定理表明, 沿着 (P) 的最优过程, 存在伴随系统 (7.82) 的解属于空间 $BV\big(0,T;H^{-1}(\Omega)\big)$ (即在 $[0,T]$ 上有界变差的取值于 $H^{-1}(\Omega)$ 的函数空间) 且满足其他条件. 为此需要追加一个假设.

(**H7**) 椭圆算子 A 的变系数 $a_i(x)$ $(i=0,\cdots,n)$ 满足条件

$$a_0(x)\geqslant 0,\quad a_0(x)-\sum_{i=1}^{n}\frac{\partial a_i(x)}{\partial x_i}\geqslant 0,\quad \forall x\in\Omega.$$

定理 7.50(最差扰动的必要条件) 设 $(\bar u,\bar w,\bar y)$ 为原极小极大问题 (P) 在假设 (H1)~(H5) 和 (H7) 下的最优三元组, 再设规范条件 (CQ1) 成立. 则存在满足 $\operatorname{supp}\lambda_1\subset Q_{ab}$ 的 $\lambda_1\in L^\infty(Q)^*$ 和伴随系统 (7.82) 的轨道 $p\in BV\big(0,T;H^{-1}(\Omega)\big)\cap L^2\big(0,T;H_0^1(\Omega)\big)\cap L^\infty(0,T;X)$ 满足

$$\int_Q\left[\Big(B^*p_1+\frac{\partial f}{\partial w}(x,t,\bar w)\Big)(w-\bar w)\right]\mathrm{d}x\,\mathrm{d}t\leqslant 0,\quad \forall w\in W_{ad}. \tag{7.83}$$

证明 该证明是通过对定理 7.41 中逼近问题 $(\mathrm{P}_{1\varepsilon})$ 的必要最优条件取极限来实现的. 令 $p_{1\varepsilon}$ 为定理 7.41 中伴随系统 (7.78) 对应于 $(u_\varepsilon,y_{1\varepsilon})$ 的强解. 对系统中两部分都乘上 $v\in W_0^{2,1,\infty}(Q)$ 并分部积分, 则得积分恒等式

$$\begin{aligned}&\int_Q p_{1\varepsilon}\Big(\frac{\partial v}{\partial t}+Av\Big)\mathrm{d}x\,\mathrm{d}t\\ =&\int_Q\frac{\partial g}{\partial y}(x,t,y_{1\varepsilon}+\bar y_2)v\,\mathrm{d}x\,\mathrm{d}t\\ &-\int_Q 2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar y_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar y_2)v\,\mathrm{d}x\,\mathrm{d}t\quad(\text{当}v\in W_0^{2,1,\infty}(Q)).\end{aligned}$$

(7.78) 的强解 $p_{1\varepsilon}$ 可表为形式

$$p_{1\varepsilon}(t)=-\int_t^T S^*(\tau-t)\Big(\frac{\partial g}{\partial y}(\tau,x,y_{1\varepsilon}+\bar y_2)-2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar y_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar y_2)\Big)\mathrm{d}\tau,$$

这对任意 $t\in[0,T]$ 成立, 其中 $S^*(\cdot)$ 是 $-A^*$ 生成的强连续半群. 由 Brézis 和 Strauss[177] 的结果知, (H7) 保证了 $S^*(\cdot)$ 在 $L^1(\Omega)$ 中的压缩性质. 由此以及引理 7.48 的第一个估计, 可找到与 ε 和 t 无关的常数 $M>0$ 使得

$$\begin{aligned}&\|p_{1\varepsilon}(t)\|_{L^1(\Omega)}\\ \leqslant&\int_t^T\Big\|S^*(\tau-t)\Big(\frac{\partial g}{\partial y}(x,\tau,y_{1\varepsilon}+\bar y_2)-2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar y_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar y_2)\Big)\Big\|_{L^1(\Omega)}\mathrm{d}\tau\\ \leqslant&\int_t^T\|S^*(\tau-t)\|\cdot\Big\|\frac{\partial g}{\partial y}(x,\tau,y_{1\varepsilon}+\bar y_2)-2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar y_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar y_2)\Big\|_{L^1(\Omega)}\mathrm{d}\tau\\ \leqslant&\Big\|\frac{\partial g}{\partial y}\Big\|_1+\|2\varepsilon\alpha_\varepsilon'(y_{1\varepsilon}+\bar y_2)\alpha_\varepsilon(y_{1\varepsilon}+\bar y_2)\|_1\leqslant M<\infty\end{aligned}$$

对任意 $t \in [0,T]$ 和 $\varepsilon > 0$ 成立. 这说明 $\{p_{1\varepsilon}\}_{\varepsilon>0}$ 在 $\mathcal{C}\big([0,T];L^1(\Omega)\big)$ 中是有界的. 另外, 由 (7.78) 和引理 7.48 得 $\left\{\partial p_{1\varepsilon}/\partial t - A^* p_{1\varepsilon}\right\}_{\varepsilon>0}$ 在 $L^1(Q)$ 中有界. 则 Sobolev 嵌入定理保证了 $\left\{\partial p_{1\varepsilon}/\partial t\right\}$ 在 $L^1\big(0,T;H^{-1}(\Omega)\big)$ 中的有界性. 进一步, 基于 (7.78) 和前面的估计, 也可得 $\{p_{1\varepsilon}\}$ 在 $L^2\big(0,T;H_0^1(\Omega)\big)$ 和 $L^\infty(0,T;X)$ 中的有界性. 现在由标准的紧性推理和 $L^\infty(0,T;X)$ 对偶于一个可分 Banach 空间这个事实, 即可找到

$$p_1 \in BV\big(0,T;H^{-1}(\Omega)\big) \cap L^2\big(0,T;H_0^1(\Omega)\big) \cap L^\infty(0,T;X)$$

和 $\{p_{1\varepsilon}\}$ 的一个子序列 (不再重记下标) 满足

$$\begin{aligned}
&p_{1\varepsilon}(t) \to p_1(t) \quad \big(H^{-1}(\Omega) \text{ 中的强收敛}\big),\\
&p_{1\varepsilon} \to p_1 \quad \big(L^2\big(0,T;H_0^1(\Omega)\big) \text{ 中的强收敛}\big),\\
&p_{1\varepsilon} \to p_1 \quad \big(L^\infty(0,T;X) \text{ 中的弱 * 收敛}\big),
\end{aligned}$$

其中 $\varepsilon \downarrow 0$. 在上面的积分恒等式中在 $\varepsilon \to 0$ 时取极限并利用 (序列版本的) 引理 7.49, 即得 p_1 满足伴随系统 (7.82). 最后应用定理 7.40 中的收敛结果和 $L^2(0,T;X)$ 中的强收敛 $p_{1\varepsilon} \to p_1$, 对定理 7.41 中的对应条件在 $\varepsilon \downarrow 0$ 时取极限就得到必要条件 (7.83). △

推论 7.51(最差扰动的砰–砰关系) 在定理 7.50 的假设下, 有

$$\bar{w}(x,t) = c \quad \text{和} \quad \bar{w}(x,t) = d$$

分别在集合

$$\left\{(x,t) \in Q \mid \big(B^* p_1\big)(x,t) + \frac{\partial f}{\partial w}\big(x,t,\bar{w}(x,t)\big) < 0\right\}$$

和

$$\left\{(x,t) \in Q \mid \big(B^* p_1\big)(x,t) + \frac{\partial f}{\partial w}\big(x,t,\bar{w}(x,t)\big) > 0\right\}$$

上几乎处处成立, 其中 $p_1(x,t)$ 是伴随系统 (7.82) 的对应解.

证明 由 (7.83) 易得. △

通过对逼近问题 $(\mathrm{P}_{2\varepsilon})$ 必要最优条件取极限, 下面推导原极小极大问题 (P) 中 Dirichlet 边界控制的必要最优条件. 对此的关键是证明 (7.72) 的从 $L^\infty(\Sigma)$ 到 $L^\infty(\Omega)$ 的适度解算子 $\mathcal{L}$ 是连续的. 下述定理保证该性质并建立了原状态约束问题中的 Dirichlet 边界控制的必要最优条件.

定理 7.52(Dirichlet 边界控制的必要最优条件) 设 $(\bar{u},\bar{w},\bar{y})$ 为极小极大问题 (P) 在假设 (H1)~(H4) 和 (H6) 下的最优三元组. 再假设规范条件 (CQ2) 成立. 则

存在满足 $\operatorname{supp}\lambda_2 \subset Q_{ab}$ 的测度 $\lambda_2 \in L^\infty(Q)^*$ 使得

$$\begin{aligned}0 \leqslant &\int_\Sigma \left[\mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y})\Big) + \frac{\partial h}{\partial u}(s,t,\bar{u})\right](u-\bar{u})\,\mathrm{d}s\,\mathrm{d}t \\ &+ \int_\Sigma (u-\bar{u})\big(\mathcal{L}^*\lambda_2\big)(\mathrm{d}s\,\mathrm{d}t) \quad \forall u \in U_{ad}.\end{aligned} \tag{7.84}$$

证明 取 $(u_\varepsilon, y_{2\varepsilon})$ 为问题 (P_ε) 的最优解, 根据定理 7.44 可使其在 $\varepsilon \downarrow 0$ 时强收敛到给定最优解 $(\bar{u}, \bar{y}_2)$, 并对任意 $\varepsilon > 0$ 满足定理 7.45 的必要最优条件. 由定理 7.45 直接可得

$$\begin{aligned}0 \leqslant &\int_\Sigma \left[\mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y}_1+y_{2\varepsilon}) + 2\varepsilon\alpha'_\varepsilon(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\Big)\right. \\ &\left.+\frac{\partial h}{\partial u}(s,t,u_\varepsilon)\right](u-u_\varepsilon)\,\mathrm{d}s\,\mathrm{d}t \\ &+2p\int_0^T \|u_\varepsilon-\bar{u}\|_U^{p-2}\Big(\int_\Gamma (u_\varepsilon-\bar{u})(u-u_\varepsilon)\,\mathrm{d}s\Big)\,\mathrm{d}t, \quad \forall u \in U_{ad}.\end{aligned} \tag{7.85}$$

现在需要阐明沿 $\varepsilon \downarrow 0$ 的一个子网在 (7.85) 中取极限的问题. 定理 7.44 中的收敛性结果和算子 $\mathcal{L}^*\colon L^2\big(0,T;L^2(\Omega)\big) \to L^2\big(0,T;L^2(\Gamma)\big)$ 的连续性 (见 Lasiecka 和 Triggiani [743]) 蕴涵

$$\begin{aligned}&\int_\Sigma \left[\mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y}_1+y_{2\varepsilon})\Big) + \frac{\partial h}{\partial u}(s,t,u_\varepsilon)\right](u-u_\varepsilon)\,\mathrm{d}s\,\mathrm{d}t \\ \to &\int_\Sigma \left[\mathcal{L}^*\Big(\frac{\partial g}{\partial y}(x,t,\bar{y})\Big) + \frac{\partial h}{\partial u}(s,t,\bar{u})\right](u-\bar{u})\,\mathrm{d}s\,\mathrm{d}t, \quad \forall u \in U_{ad},\end{aligned}$$

并且 (7.85) 中的最后一项在 $\varepsilon \downarrow 0$ 是收敛到零. 为从 (7.85) 得到 (7.84), 余下需要证明

$$\begin{aligned}&\int_\Sigma (u-u_\varepsilon)\mathcal{L}^*\Big(2\varepsilon\alpha'_\varepsilon(\bar{y}_1+y_{2\varepsilon})\alpha_\varepsilon(\bar{y}_1+y_{2\varepsilon})\Big)\,\mathrm{d}s\,\mathrm{d}t \\ \to &\int_\Sigma (u-\bar{u})\big(\mathcal{L}^*\lambda_2\big)(\mathrm{d}s\,\mathrm{d}t), \quad (\text{当} u \in U_{ad})\end{aligned}$$

沿 $\varepsilon \downarrow 0$ 的某子网成立. 考虑到引理 7.49, 这显然源于伴随算子

$$\mathcal{L}^*\colon L^\infty(Q)^* \to L^\infty(\Sigma)^*$$

的弱 * 连续性. 而此连续性是适度解算子 $\mathcal{L}$ (作为从 $L^\infty(\Sigma)$ 到 $L^\infty(Q)$ 的算子) 的强连续性的直接推论. 为证 $\mathcal{L}$ 的该连续性, 考虑抛物方程的广义解理论和前面的讨论如下.

取定 (7.70) 中 Dirichlet 边界条件中的函数 $v \in L^2(\Sigma)$. 应用 Lions 的书 [791] 中的定理 9.1, 存在唯一的 $y(v) \in L^2(Q)$ (称为 (7.70) 的广义解), 使得

$$\int_Q y(v)\Big(-\frac{\partial z}{\partial t} + A^* z\Big)\,\mathrm{d}x\,\mathrm{d}t = -\int_\Sigma v\frac{\partial v}{\partial \nu_A}\,\mathrm{d}s\,\mathrm{d}t \tag{7.86}$$

在 $z \in H^{2,1}(Q)$ 且满足 $z(s,t) = 0$ (当 $(s,t) \in \Sigma$ 和 $z(T,x) = 0$) 时成立. 取系统 (7.70) 由 $v \in L^\infty(\Sigma)$ 生成的适度解 $y = \mathcal{L}v$, 下证该 y 是 (7.86) 意义下 (7.70) 的广义解.

为此把给定的 Dirichlet 边界控制 v 看成空间 $L^p(0,T;U)$ (p 足够大) 的元并注意到定义域空间 $\mathcal{D}(\Sigma)$ 在 $L^p(0,T;U)$ 中稠密, 即存在控制序列 $\{v_k\} \subset \mathcal{D}(\Sigma)$ 满足

$$v_k \to v \quad (k \to \infty, L^p(0,T;U) \text{ 中的强收敛}).$$

众所周知, 对任意 $v_k \in \mathcal{D}(\Sigma)$, 系统 (7.70) 容许唯一的经典解 y_k, 自动是该系统的适度解和广义解. 因此 $y_k = \mathcal{L}v_k$ 且 y_k 满足 (7.86) $(\forall k \in \mathbb{N})$. 进一步, 由定理 7.33 中的正则性结果得

$$\big\|\mathcal{L}v - y_k\big\|_{\mathcal{C}([0,T];X)} = \big\|\mathcal{L}v - \mathcal{L}v_k\big\|_{\mathcal{C}([0,T];X)} \to 0 \quad (k \to \infty).$$

综合所有这些事实就得到估计

$$\begin{aligned}
&\left|\int_Q \mathcal{L}v\Big(-\frac{\partial z}{\partial t} + A^* z\Big)\,\mathrm{d}x\,\mathrm{d}t + \int_\Sigma v\frac{\partial z}{\partial \nu_A}\,\mathrm{d}s\,\mathrm{d}t\right| \\
\leqslant &\left|\int_Q (\mathcal{L}v - y_k)\Big(-\frac{\partial z}{\partial t} + A^* z\Big)\,\mathrm{d}x\,\mathrm{d}t\right| + \left|\int_\Sigma (v - v_k)\frac{\partial z}{\partial \nu_A}\,\mathrm{d}s\,\mathrm{d}t\right| \\
\leqslant &\big\|\mathcal{L}v - y_k\big\|_{\mathcal{C}([0,T];X)} \cdot \Big\|-\frac{\partial z}{\partial t} + A^* z\Big\|_{L^2(0,T;X)} T^{1/2} \\
&+ \big\|v - v_k\big\|_{L^p(0,T;U)}\Big\|\frac{\partial z}{\partial \nu_A}\Big\|_{L^2(0,T;U)} T^{1/\widetilde{q}} \to 0 \quad (k \to \infty),
\end{aligned}$$

其中 $\widetilde{q} := 2(p-1)/p-2$. 这给出

$$\int_Q \mathcal{L}v\Big(-\frac{\partial z}{\partial t} + A^* z\Big)\,\mathrm{d}x\,\mathrm{d}t = -\int_\Sigma v\frac{\partial z}{\partial \nu_A}\,\mathrm{d}s\,\mathrm{d}t$$

在 $z \in H^{2,1}(Q)$ 且满足 $z(s,t) = 0$ (当 $(s,t) \in \Sigma$ 和 $z(T,x) = 0$) 时成立. 该式意味着对任意 $v \in L^\infty(\Sigma)$, $y = \mathcal{L}v$ 是 (7.70) 的一个广义解. 最后由广义解的唯一性和广义解算子 (从 $L^\infty(\Sigma)$ 到 $L^\infty(Q)$) 的连续性 (见前面提到的 Lions 的书), 得所论的线性算子 $\mathcal{L}$ (从 $L^\infty(\Sigma)$ 到 $L^\infty(Q)$) 是连续的. 证毕. △

总结所得结果即得下述定理, 它提供了原极小极大问题的最差扰动和 Dirichlet 最优控制的必要条件.

定理 7.53(刻画极小极大最优解) 设 $(\bar{u},\bar{w})$ 为极小极大问题 (P) 的最优解, $\bar{y}$ 为抛物系统 (7.61) 对应的轨道. 假设 (H1)~(H7) 皆成立且有约束规范条件 (CQ1) 和 (CQ2). 则存在满足 $\operatorname{supp}\lambda_i \subset Q_{ab}$ 的 $\lambda_i \in L^\infty(Q)^*$ $(i=1,2)$ 和 (7.82) 的伴随轨道

$$p_1 \in BV\big(0,T;H^{-1}(\Omega)\big) \cap L^2\big(0,T;H_0^1(\Omega)\big) \cap L^\infty(0,T;X)$$

使得最优条件 (7.83) 和 (7.84) 成立.

注 7.54(反馈控制设计) 本节建立的结果可用来确定开环极小极大控制问题 (P) 的最差扰动和最优边界控制的结构. 它们对闭环抛物控制系统的极小极大设计也是有用的, 此时的目的是构造一种反馈控制, 它依赖于状态变量并保证在容许区域 W_{ad} 中的任意扰动下的满意 (至少是稳定) 表现, 并在最差扰动时达到最佳. 其中反馈控制问题已在 Mordukhovich [905, 918] 和 Mordukhovich 与 Shvartsman [957] 中得到研究, 读者可在这些文章里找到更多的讨论和文献. 这些文章研究的一个代表性问题是

$$\min J(u) := \max_{w\in W_{ad}} \int_0^T \Big|u\big(y(x_0,t)\big)\Big|\,\mathrm{d}t$$

其中要求 $u \in U_{ad}$ 且满足抛物系统

$$\begin{cases} \dfrac{\partial y}{\partial t} + Ay = w(t) \quad (\text{在 } Q \text{ 中几乎处处成立}), \\ y(x,0) = 0, \quad x \in \Omega, \\ y(s,t) = u(t), \quad t \in \Sigma, \end{cases}$$

逐点状态约束

$$\big|y(x_0,t)\big| \leqslant \eta, \quad \forall t \in [0,T]$$

和反馈控制律

$$u(t) = u\big(y(x_0,t)\big),$$

其中 $x_0 \in \Omega$ 为给定点 (在该点收集所有有关系统输出的信息), 容许扰动和控制区域分别定义于 (7.63) 和 (7.64), 且 A 为自伴强一致椭圆算子, 定义为

$$A := -\sum_{i,j=1}^n a_{ij}(x)\frac{\partial^2}{\partial x_i \partial x_j} - c,$$

这里 $c \in \mathbb{R}$, $a_{ij} \in \mathcal{C}^\infty(\operatorname{cl}\Omega)$. 除了探讨有效的开环变分分析和逼近过程, 这些文章还涉猎了抛物动态的单调性质和轨道在无穷水平线上的渐近特性, 这可用来发展反馈次最优控制的有效的极小极大设计, 并在大范围内保证极度非线性闭环控制系统中需要的稳定性; 更多的细节, 数值分析, 以及尚待解决的问题请见提到的这些文章.

7.5　第 7 章的评注

7.5.1　分布与集总 (集中) 参数控制系统

第 7 章专门研究了具有所谓“分布参数”的某些类型的动态优化和最优控制系统. 传统上控制理论分为两种, 一是常微分方程控制的系统 (及其离散时间版本), 一是动态由更复杂的方程 (比如涉及各种时滞, 或更一般地, 泛函微分方程) 描述的系统, 或是由各种偏微分方程 (椭圆、抛物、双曲等) 控制的系统. 该分类主要依据所论状态空间的维数 (有穷或无穷), 这至少从最优控制的早期就开始了. 特别地, 常微分方程控制的系统

$$\dot{x}(t) = f\big(x(t), u(t), t\big), \quad x(t) \in \mathbb{R}^n, \quad t \in [a, b] \tag{7.87}$$

的自然状态空间是有穷维的. 其当前状态, 即 (7.87) 中的 $x(t)$ $(t \in [a, b])$, 对给定的控制 $u(t)$, 可由给定的初始状态向量 $x(a) = x_0 \in \mathbb{R}^n$ 完全确定. 这种有穷维系统也称为具有“集总参数”(亦译为“集中参数”) 的系统.

与此相反, 即使对最简单的时滞控制系统 (具有状态变量上的单一时滞 $\theta > 0$)

$$\dot{x}(t) = f\big(x(t), x(t-\theta), u(t), t\big), \quad x(t) \in \mathbb{R}^n, \quad t \in [a, b], \tag{7.88}$$

其自然状态空间也是无穷维的. 这是因为需要知道初始函数 $x(t)$ 在整个“初始尾”区间 $t \in [a-\theta, a]$ 的值来决定当前状态 $x(t)$ $(t \in [a, b])$. 类似的 (经常更复杂的) 情形也出现在由偏微分方程、积分方程等控制的控制过程中.

粗略地讲, 集总参数系统是指由有穷维空间中 (7.87) 类型的常微分方程 (及其离散时间的版本, 和更一般的微分与离散包含模型) 描述的控制过程, 而分布参数系统的描述则涉及各种无穷维状态空间.

这种传统的大面上的分类似乎是很有局限的, 不能反映出状态动态系统的很多不能仅以状态空间的维数来刻画的特殊性质. 特别地, 第 6 章研究的动态优化和控制问题具有无穷维的状态空间, 根据此分类无疑属于分布参数的; 但另一方面, 所涉及的动态的常微分形式却在所发展的方法和所建立的结果中起着至关重要的作用, 尽管状态空间的无穷维性质的确需要更复杂的办法.

第 7 章涉及几类分布参数的发展控制系统, 它们的动态描述互相差别很大, 也与第 6 章中的常微分情形截然不同. 除了变分分析的一般技术和结构 (这对第 6 章和第 7 章的主要研究基本是一样的), 本章对分布系统建立的结果和方法本质上是基于所论系统的特殊性质.

7.5.2 状态变量具有时滞的系统

7.1 节致力于具有时滞的动态优化问题. 这种系统 (也称为时滞/差、遗传、推迟、微分–差分、泛函–微分系统, 或具有后效或自变量偏差的系统等) 中的动态过程依赖于"史前"的状态, 从而使之成为无穷维的, 即使在有穷维状态变量 $x \in \mathbb{R}^n$ 的情形亦如此. 遗传系统 (没有控制) 的定性研究起始于 Volterra (见他的书 [1298]), 而密集的发展开始于 20 世纪 50 年代 Myshkis 的开创性的书 [989]; 见后来的书, Bellman 与 Cooke [94], Élsgolts 与 Norkin [404], Hale [538], Kolmanovskii 与 Myshkis [694], 及其所录文献.

时滞最优控制最早的结果是由 Kharatishvili [678] 得到的一个 Pontryagin 最大值原理的类似版本, 所研究的系统具有 (7.88) 类型的时滞微分方程, 涉及状态变量的单一时滞. 接着这些结果被推广到状态和控制变量具有变动的和分布的时滞的系统, 系统具有各种约束, 包括本质上是无穷维的类型等. 对时滞微分系统的最优控制的早期贡献, 这里提一下这些研究工作: Banks [78], Friedman [476], Gabasov 与 Churakova [483], Gabasov 与 Kirillova [486], Halanay [537], Krasovskii [700], Oğuztöreli [1018] 和 Warga [1315]. 在该方向及相关方向上, 对状态和控制变量具有各种时滞并且带有更一般的约束的控制遗传的进一步结果在此之后在诸多文献中有探讨, 如可见文献 [81, 101, 275, 281, 485, 486, 506, 679, 694, 696, 701, 867, 901, 1015, 1019, 1173, 1174, 1321, 1323].

7.5.3 中立型遗传系统

与常微分系统 (7.87) 及其时滞微分版本 (7.88) 皆本质不同, 有一类很有趣的遗传系统, 其描述为

$$\dot{x}(t) = f\big(x(t), x(t-\theta), \dot{x}(t-\theta), u(t), t\big). \tag{7.89}$$

它称为"中立型泛函微分系统", 也简称为"中立型控制系统". 在具有后效性的微分方程定性理论的研究中, Élsgolts [403] 首先在"中立型自变量具有偏差的微分方程"的名字下讨论了这类系统 (不具有控制的情形); 也可见 Bellman 与 Cooke [94], Hale [538]. 系统 (7.88) 和 (7.89) 的主要不同在于, 后者在速度变量上有时滞, 而不仅仅在状态和/或控制变量上. 这造成了中立系统 (7.89) 上的控制问题比对应于 (7.87) 和 (7.88) 的问题要复杂得多. 特别地, 在一般的非凸情形下, 中立型控制系统没有 Pontryagin 最大值原理的类似结果; 见注 6.41 中的讨论及例子, 它取自 Gabasov 与 Kirillova [485].

中立型变分和控制系统必要最优条件的最早结果似乎是分别在 Hughes [586], 和 Kamenskii 与 Hvilon [663] 中独立得出的; 也可见早期的文章 Sabbagh [1185] 和 Kent [668]. 各种后续研究可见 Angell 与 Kirsch [19], Banks 与 Kent [79], Banks 与

Manitius [81], Chukwu [241], Élsgolts 与 Norkin [404], Gabasov 与 Kirillova [485], Gorelik 与 Mordukhovich [513, 514], Gusakova [529], Jacobs 与 Langenhop [626], Kisielewicz [682], Kharatishvili 与 Tadumadze [679], Kolmanovskii 与 Nosov [695], Kolmanovskii 与 Shaikhet [696], Mansimov [843], Melikov [868], Mordukhovich [895, 896, 901], Mordukhovich 与 Sasonkin [945], Salamon [1187], Tadumadze 与 Alkhazishvili [1242], 及其所录的文献.

除了前面提到的 Pontryagin 最大值原理对非凸中立型系统不再成立, 还有其他重要的问题使得中立型系统因为速度变量具有时滞而与其常微分和时滞微分版本非常不同; 这特别包括如下问题: 一阶必要最优条件中的各种伴随系统及相关课题 [81, 485, 513, 529, 663, 668, 679, 868, 901]; 最优条件中不可避免的跳跃; 保证松弛稳定性的条件上的更多限制 [682], 对一些时间变化特性结果的本质依赖, 这些特性是关于系统中涉及速度时滞的那些分量的 [81, 485, 695, 901, 945]; 不连续初始条件对可控性、可观测性、最优性、对偶等方面主要结果的形式的影响 [895, 896, 901, 1242]; 对 $\dot{x}(t-\theta)$ 非线性的中立型系统一些中间类型 (介于一阶和二阶) 必要最优条件的出现 [513, 901]; 控制变量不具有约束的非线性中立型系统的 Legendre-Clebsch 类型的新二阶条件 [514, 901]; 奇异控制的各种必要最优条件 [514, 843, 901] 等. 另外, 与常微分和时滞微分系统不同, 对不具有端点约束的光滑中立型系统的有限差分逼近一般不成立所谓的“相近 (approximative) 最大值原理”; 见第 6 章的例 6.70.

可以看到, 中立型系统表现出很多和离散最优控制系统的共同点. 在某种意义上这似乎并不奇怪, 因为 (7.89) 可以看成关于速度 $\dot{x}(t)$ 的离散系统, 其中时滞 $\theta > 0$ 充当了离散步长的角色. 另一方面, 对中立型系统得到的某些结果在双曲类型偏微分方程的控制理论中有相似的版本, 特别是对所谓的 Goursat-Darboux 系统; 请比照 Ashchepkov 与 Vasiliev [42], Cernea [233], Gavrilov 与 M. Sumin [500], Mahmudov [827], Mansimov [844], Plotnikov 与 V. Sumin [1085], Srochko [1221], 和 Vasiliev [1281]. 与此同时, 某些双曲系统 (比如那些由电报或类似方程控制的系统) 可以等价地约化为中立型系统; 见 Kolmanovskii 与 Nosov [695] 及所录文献.

7.5.4　时滞微分包含

时滞系统动态优化中比 (7.88) 更复杂的一些问题可由如下类型的“时滞微分包含”来描述:

$$\dot{x}(t) \in F\big(x(t), x(t-\theta), t\big), \quad \text{a.e. } t \in [a, b]. \tag{7.90}$$

它具有初始和端点条件

$$x(t) = c(t), \quad t \in [a-\theta, a), \quad \big(x(a), x(b)\big) \in \Omega.$$

该类问题首先以名字“微分-差分包含”见于 Clarke 与 Watkins [274]. 在 $F(x, y, t)$ 的紧性和凸性, 以及 $F(\cdot, \cdot, t)$ 的 Lipschitz 连续性假设下, 他们对 Mayer 问题, 即在

(7.90) 的绝对连续轨道中极小化 $\varphi(x(b))$, 建立了必要最优条件. 这些条件以 Clarke 的 Hamilton 函数 ([255] 中函数的推广) 形式给出, 其中的横截条件由 Clarke 自己的广义微分结构 $\partial_C\varphi$ 和 $N_C(\cdot;\Omega)$ 给出. 除了必要最优条件, [274] 还包括值函数广义梯度的计算 (它依赖于端点扰动) 及这些结果在局部可控性上的应用.

[274] 中的 Hamilton 函数条件在 Clarke 与 Wolenski [276] 中被推广到了涉及如下更一般遗传包含的 Bolza 问题:

$$\dot{x}(t) \in F(x_t, t), \quad \text{a.e. } t \in [a, b],$$

其中 $x_t\colon [-r, 0] \to \mathbb{R}^n$ 由 $x_t(s) := x(t+s)$ 给出. 文献 [276] 中的必要条件是通过无穷维空间中近邻分析的扰动技术得到的.

具有形式 (7.90) 的时滞微分包含优化问题的另一方法见 Minchenko [878], 该文把 Polovinkin 与 Smirnov[1094], 和 Frankowska[465] 的原空间构造推广到了时滞系统的情形. 文献 [878] 中得到的必要最优条件由"切向逼近"给出, 一般来说, 和文献 [274] 中结果是独立的. 该方向的进一步结果可见 Minchenko 与 Volosevich [881]. 这里还向读者指出 Cernea 的近期文章 [234], 其中以切向技术和方向导数建立了一些时滞微分包含的二阶必要最优条件.

研究由时滞微分包含 (7.90) 控制的优化问题的"离散逼近"方法是在 Mordukhovich [921] 和 Mordukhovich 与 Trubnik [959] 中建立的. 文献 [915] 中的结果阐明了离散逼近系统 (7.90) 的适定性和强收敛过程, 并建立了广义 Euler-Lagrange 和 Hamilton 函数形式 (后者只对凸值包含成立) 的必要最优条件, 用的方法是基本法锥和次梯度的部分凸化, 这和文献 [274, 276] 中对 Hamilton 包含的完全凸化是不同的. 进一步的关 Bolza 问题的结果见 Mordukhovich 与 L. Wang [973], 其中的容许轨道取自满足如下多值"初始尾"条件

$$x(t) \in C(t), \quad \text{a.e. } t \in [a - \theta, a)$$

的时滞微分包含 (7.90). 该条件是时滞系统特有的, 提供了控制和优化的更多来源.

7.5.5 中立型微分包含

对由中立型泛函微分包含

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[x(t) - Ax(t-\theta)\Big] \in F\big(x(t), x(t-\theta), t\big), \quad t \in [a, b], \quad \text{a.e.} \tag{7.91}$$

控制的优化问题, 其必要最优条件首先由 Mordukhovich 与 L. Wang 建立, 其中 Mayer 问题的讨论见文献 [972], 具有端点约束的 Bolza 问题的详尽研究见文献 [974]; 另外文献 [977] 中讨论了更一般和复杂的非自治系统的情形, 其中 (7.91) 中的矩阵 $A = A(t)$ 依赖于时间变量.

注意 (7.91) 中左边的中立型算子以所谓的“Hale 形式”[538] 给出, 对文献 [972, 974, 977] 中的技术与结果这是至关重要的. 这些文章中的方法是源于 Mordukhovich[915], 基于“离散逼近”. 当然, 中立系统的情形明显地更加复杂; 与常微分和时滞微分包含相比, 没有“松弛稳定性”则得不到任何结果. 而该稳定性要求对非凸中立型系统有时是很苛刻的; 其成立的某些充分条件可见 Kisielewicz [682]. 文献 [972, 974, 977] 中的主要必要最优条件以广义 Euler-Lagrange 形式给出, 这蕴涵着对应的 Weierstrass-Pontryagin 最大值条件的类似版本, 也蕴涵 (应用 Rockafellar 的对偶化定理 [1162]) 改进了的 Hamilton 条件 (这需要所设的松弛稳定性, 特别是对中立型的凸值泛函微分包含).

Ortiz 最近的文章 [1021] 涉及了具有变化时滞 $\theta = \theta(t)$ 的中立型系统的如下推广 Bolza 问题的必要条件:

$$\min \varphi\big(x(a), x(b)\big) + \int_a^b \vartheta\big(x(t), x(t-\theta), \dot{x}(t-\theta), t\big)\,\mathrm{d}t, \tag{7.92}$$

其中函数 φ 和 ϑ 都可以是增广实值的. 该文是对 Ortiz 与 Wolenski [1022] 的扩充. 文献 [1022] 讨论的是状态中具有时滞 (不是中立) 的推广 Bolza 问题, 其中 (7.92) 中的被积项 ϑ 不依赖于时滞速度项 $\dot{x}(t-\theta)$ 而是依赖于时滞状态变量 $x(t-\theta)$. 文献 [1021] 中的主要假设是被积项 ϑ 关于 $\dot{x}(t)$ 和 $\dot{x}(t-\theta)$ 的两个速度变量的二元联合凸性, 其方法基于“解耦技术”, 它由 Clarke[258] 对非时滞 Bolza 问题提出, 然后由 Ortiz 与 Wolenski[1022] 在时滞系统中建立. 可以看到, 文献 [1021], [258], [1022] 中结果的形式都是 Clarke 类型的全凸化 Euler-Lagrange 和 Hamilton 包含, 而不是 Mordukhovich 与 L. Wang [974] 中大幅改进的涉及基本法锥和次微分部分凸化的形式, 后者推广了 Mordukhovich 与 Rockafellar 对非时滞系统的一些形式; 比照第 6 章.

7.5.6　微分–代数系统

7.1 节专门研究了这样一类动态优化问题, 其动态约束由如下相互关联的时滞微分包含和线性时滞–代数方程描述:

$$\begin{cases} \dot{z}(t) \in F\big(x(t), x(t-\theta), z(t), t\big), & \text{a.e. } t \in [a,b], \\ z(t) = x(t) + Ax(t-\theta), & t \in [a,b]. \end{cases} \tag{7.93}$$

这是具有分布参数的一类新的最优控制问题. 它一方面可作为扩充中立型包含的变分问题, 另一方面它相关于一类特殊的时滞微分–代数系统, 这种系统受控于一般的时滞微分包含, 其中的“慢”和“快”变量之间具有线性时滞代数联系. 注意 7.1 节考虑的问题 (DA) 的 Bolza 泛函中的被积项 ϑ 依赖于慢和快两个变量 (分别记为 z 和 x), 也依赖于慢变量的时间导数, 而快变量对时间可以不可导.

系统 (7.93) 可以写成形式

$$\frac{\mathrm{d}}{\mathrm{d}t}\Big[x(t)+Ax(t-\theta)\Big]\in F\big(x(t),x(t-\theta),x(t)+Ax(t-\theta),t\big),\quad \text{a.e. } t\in[a,b].$$

这是 (7.91) 的一个推广, 可以化归为一般的中立型包含

$$\dot{x}(t)\in G\big(x(t),x(t-\theta),\dot{x}(t-\theta),t\big),\quad \text{a.e. } t\in[a,b]. \tag{7.94}$$

这里需要假设 $x(t)$ 在 $[a,b]$ 上绝对连续. 本书没有用到该假设, 而是设 $x(t)+Ax(t-\theta)$ 绝对连续. 类似地, (DA) 中的费用函数在该代换下变成中立型 Bolza 形式 (7.92). 因此 (DA) 可以作为一般中立型包含的 Bolza 类型变分问题的一个特殊情形来处理. 但是, 如此则丢失了 (DA) 的一些主要特点, 而这些特点对 7.1 节的方法和结果都是举足轻重的. 这里所指的特点是, 动态约束 (7.94) 和费用函数 (7.92) 都不是依赖于 $\dot{x}(t)$ 和 $\dot{x}(t-\theta)$, 而是同样的线性组合 $x(t)+Ax(t-\theta)$ 的导数. 正是这个原因书中将此线性组合看成 (7.93) 中的新状态变量并以其自然形式研究 (DA), 这同时强调了系统动态上的时滞微分和线性的代数约束.

在 $\theta=0$ 时的非时滞情形, 系统 (7.93) 是所谓的控制“微分–代数方程”(简称 DAE) 的包含扩充. 这种方程出现在很多实际应用中, 特别包括过程系统工程、机器人、具有完整和非完整约束的力学系统和等; 见 Brennan, Campbell, 与 Pretzold 的书 [174], 那里有大量的例子、讨论和文献. DAE 控制系统一般具有形式

$$\begin{cases}\dot{z}=f\big(z(t),x(t),u(t),t\big), & \text{a.e. } t\in[a,b],\\ 0=g\big(z(t),x(t),u(t),t\big), & \text{a.e. } t\in[a,b].\end{cases} \tag{7.95}$$

它与其他一些不同名称下的特殊控制系统紧密相关, 这包括隐含系统、奇异系统、描述系统 (descriptor) 等; 如可见 Dai [305], Devdariani 与 Ledyaev [327] 及所录文献. 在早期俄文文献中, 该系统以“非关于导数而解的控制系统”这个长名字来研究的, 比如可见 Vasiliev [1279], Gabasov 与 Kirillova [485], Gusakova [528], Kurina [730], Mordukhovich [901]. 注意 (7.95) 中的 DAE 可看成奇异扰动的控制系统

$$\begin{cases}\dot{z}=f\big(z(t),x(t),u(t),t\big), & \text{a.e. } t\in[a,b],\\ \varepsilon\dot{x}(t)=g\big(z(t),x(t),u(t),t\big), & \text{a.e. } t\in[a,b]\end{cases} \tag{7.96}$$

在 $\varepsilon\downarrow 0$ 时的极限情形. 但众所周知, (7.96) 的最优解在 $\varepsilon\downarrow 0$ 时取极限的过程需要很苛刻的假设; 请比照 Artstein 与 Gaitsgory [30], Bensoussan [100], Dontchev 与 Zolezzi [367], Kokotović, Khalil 与 O'Reilly[693], 及所录文献.

据作者所知, 具有 DAE 动态 (7.95) 的控制系统的最先进的必要最优条件是 Pinho 与 Vinter[1079] 在“指标为一”的假设下得到的. 该文论述了 (强) Pontryagin

最大值原理对这样的系统并不成立, 并在一些凸性假设下证明了该原理. 该文还对具有非光滑微分 (不是代数) 动态和具有非光滑费用和端点约束函数的系统建立了新的"弱"最大值原理形式的必要条件. 但文献 [1079] 中关键的指标为一的假设似乎还是限制很大, 在许多具有实际意义的微分–代数控制系统中不成立; 如可见文献 [174, 1048].

7.5.7 时滞的正则化角色

7.1 节中的结果主要基于 Mordukhovich 与 L. Wang 的文章 [975, 976]. 为得到所论的微分–代数控制系统的必要最优条件, 用到了"离散逼近方法"的一个变体, 其中考虑到了时滞 $\theta > 0$. 该时滞事实上是一个正则化因子, 从而可以完全避免指标为一的假设. 和 6.1 节常微分发展系统情形的过程大体类似, 这里建立了离散逼近的适定性和强收敛, 推导出逼近差分–代数系统的必要最优条件, 然后实现了对所得离散逼近的必要最优条件取极限, 从而在松弛稳定性的假设下对原来的由时滞微分–代数系统控制的具有端点约束的 Bolza 问题得到了广义 Euler-Lagrange 和 Hamilton 形式的新的必要最优条件.

上述方法的具体实现与常微分的情形是很不同的, 技术上要复杂得多. 另一方面, 简化起见, 7.1 节假设状态向量 (x, z) 为有穷维, 这避免了无穷维空间情形的 SNC 分析法则. 另外, 与 6.1 节不同, 这里每一步都应用更方便的基本/极限法锥和次微分的分析法则而不是对应的模糊版本. 能否对时滞系统的必要最优条件在 $\theta \downarrow 0$ 时取极限而得到没有时滞的微分–代数控制系统的有意义的结果, 至今还是个开问题.

7.5.8 偏微分控制系统

第 7 章剩下的三节涉及了一些偏微分方程控制的分布参数系统的最优控制问题. 有关偏微分控制的文献浩如烟海, 因此这里只提到与本书讨论紧密相关的一些. 读者可以在如下的书及其所录文献中找到偏微分控制和应用各个方面的更多材料: Ahmed 与 Teo [4], Balakrishnan [74], Banks 与 Kunisch [80], Barbu [82], Bensoussan, Da Prato, Delfour 与 Mitter [101], Butkovsky [209], Cherkaev [237], Denkowski, Migórski与 Papageorgiou [323], A. Egorov [392], Fattorini [432], Friedman [478], Fursikov [481], Lagnese [736], Lasiecka 与 Triggiani [746], Li 与 Yong [789], Lions [791, 792], Lurie [821], Lyashko [823], Neittaaanmäki与 Tiba [997], Tiba [1255], Tröltzsch[1271], Vasiliev [1281]. 另外也可见 Burns 最近的综述文章 [206].

似乎偏微分控制的最早的文章是 1960 年的 Butkovsky 与 Lerner [211], 该文专门研究了一维热传导方程的最优控制. 后来人们认识到, 许多偏微分控制问题可以写成无穷维空间中抽象"发展系统"的形式, 但由 Pontryagin 等 [1102] 建立的处理

常微分情形的方法一般不能用来建立偏微分控制系统的最大值原理. 其中主要的限制是文献 [1102] 中用到的凸分离定理在没有额外假设的情况下不能用于无穷维的情形. 偏微分控制系统中最大值原理不再成立 (对单点目标集) 的第一个例子是 Y. Egorov 构造的, 他也导出了相当苛刻的内点条件下无穷维控制问题中的 Pontryagin 最大值原理的恰当类似版本; 见文献 [393, 394]. 另一方面, A. Egorov[391] 对抛物和双曲系统在没有内点假设时证明了最大值原理, 其中的目标/约束集由有限个等式和不等式描述. 很久之后人们才发现这些集合具有有限余维数性质; 请比照第 6 章. 对偏微分最优控制的其他早期发展 (大多是俄文文献), 推荐读者参阅 Butkovsky, A. Egorov 与 Lurie 的综述文章 [210].

在西方, 无穷维最优控制的先行工作见 Fattorini[427] 和 Balakrishnan[73], 他们首先把强连续半群理论应用到线性控制系统上. 在其他重要的早期贡献中, 这里特别指出 Conti [284], Friedman [477], Lions [791], Russel [1184], Malanowski [830] 和 Wang [1302].

对最大值原理, 最先认识到可达集和/或目标集的有穷余维数的本质重要性的是 Li 与 Yao[786], 并在他们的文章 [787] 中得到了发展. 该方向更多的发展见 Fattorini [429], 和 Li 与 Yong [788]; 对涉及半线性和拟线性偏微分方程的各类最优控制问题以及其他无穷维控制系统更多结果、讨论和文献, 也可见 Li 与 Yong 的书 [789].

7.5.9 偏微分系统的边界控制

前面的论述和指出的文献大多与分布控制作用于状态方程的偏微分控制相关, 类似于常微分动态系统. 对理论和应用都重要的偏微分问题的一个特殊的地方是有可能考虑边界控制, 即控制函数出现在边界条件中. 众所周知, 边界控制问题比分布控制要复杂得多.

边界条件主要有 Dirichlet 和 Neumann 两类, 互相极不相同. 正如本书论述的, 一般来说 Dirichlet 边界条件对对应的解算子提供了更少的正则性, 这特别见 7.2~7.4 节. 混合 (或 Robin) 类型的边界条件作为一个规律总是类似 Neumann 类型条件的基本性质.

从历史上看偏微分系统的边界控制首见于 1968 年 Fattorini 的文章 [428], 该文的主要推动力源于近似可控性上的应用, 文中用了半群算子方法. 线性抛物方程的 Dirichlet 边界控制问题的一个重大的成果由 Washburn[1324] 建立, 这基于他早先的博士论文, 而论文导师 Balakrishnan 曾就矩形考虑过一个很有启发性的例子; 见 [74]. 对各种偏微分边界控制和相关问题的进一步重要研究成果可见 Arada 与 Raymond [23, 24], Barbu [82], Barbu, Lasiecka 与 Triggiani [83], Bonnans 与 Casas [131], Bucci [183], Cârjă[223], Casas [225], Casas, Raymond 与 Zidani [226], Fattorini

[433], Fattorini 与 Murphy [435, 436], Lasiecka [739], Lasiecka 与 Triggiani [742, 743, 746], Lions [791, 792], Mordukhovich 与 Raymond [943, 944], Mordukhovich 与 Zhang [978, 979], Nowakowski 与 Nowakowska [1016], Osipov, Pandolfi 与 Maksimov [1023], Raymond [1120], Raymond 与 Zidani [1121], Tröltzsch[1271], Zuazua [1379] 等; 也请见其中的文献. 这些文章中的某些具体结果的更多参考文献和评注见 7.2~7.4 节以及下面的相关评注.

7.5.10 双曲方程的 Neumann 边界控制

本章有关偏微分方程的材料从 7.3 节中考虑的半线性波动方程的 Neumann 边界控制问题开始. 该问题并非本书中最简单的, 而是恰恰相反, 双曲方程的边界控制问题是最具挑战性的控制问题之一, 在偏微分控制中尚未得到充分的研究. 在具有逐点状态约束的情形, 似乎这种问题的第一个结果是 Mordukhovich 与 Raymond 一直到最近才在文章 [944] 中得到; 7.2 节的内容主要基于该文. 没有状态约束的边界控制问题和双曲方程控制的分布控制问题的必要最优条件的一些结果可见 Bucci [183], Fattorini [431, 432], Kasiecka 与 Triggiani [746], Lions [792], Malanowski [830], White [1328] 等及所录的文献. 但是, 双曲最优控制问题在文献中涉及不多, 没有疑问比椭圆和抛物的情形少得多. 从有穷维空间中发展方程的半群观点来看, 这种现象的一个重要的技术上的原因是, 与双曲系统相联系的由所论的对应半线性方程的无界算子生成的半群是非紧的.

本章材料的取舍主要看能否把现代变分分析的有力技术 (变分原理、无约束逼近、针形变分等) 用到所论的 Neumann 边界控制问题上, 这与第 6 章中常微分方程的情形类似. 双曲系统的讨论紧接着中立型泛函微分系统的另一原因是, 这些分布参数系统 (见 7.5.3) 的某些类型具有某些特定的相似性, 而这在以前并没有被充分利用. 注意 7.1 节中的方法涉及离散逼近技术, 这与本章中的其他方法是非常不同的.

要强调的是, 7.2 节的研究基于现代变分分析技术与双曲 Neumann 边界值问题的深刻的正则性理论的融合, 后者由 Kasiecka 与 Triggiani 在 20 世纪 80 年代末和 90 年代初建立; 见文献 [744, 745]. 偏微分方程理论的该正则性及 7.2.2 节的相关发展为后面变分技术的应用提供了必须的基础, 从而能建立状态约束双曲系统的 Neumann 边界控制的必要最优条件.

7.5.11 以 Ekeland 变分原理处理逐点状态约束

具有逐点状态约束的问题是最优控制中最难的问题之一. 值得一提的是, 对由非线性常微分方程控制的状态约束控制问题, 寻求其最大值原理成立的满意条件可能是 Dubovitskii-Milyutin 在文献 [370] 中发展的一般极点理论的基本动因; 这也可

见 Ioffe 与 Tikhomirov 的书 [618], 或 Warga 的书 [1315], 这些书涵盖了相关于常微分和泛函微分 (在文献 [1315] 中) 控制系统的状态约束问题的处理和参考文献.

起初人们通过化归为有穷维的数学规划问题来处理由椭圆和抛物方程控制的控制系统的状态约束问题. 这在偏微分文献中称为“Lagrange 乘子法”, 比如可见 Mackenroth [825]、Tröltzsch[1271] 及其中的文献.

对端点约束和 (逐点) 状态约束的常微分控制问题, 通过逼近过程来推导必要最优条件的一个强有力的方法是基于 Ekeland 变分原理的, 它始于 Ekeland 本人的文章 [397] (也可见他优秀的综述文章 [399]), 然后在 Clarke 的文章 [250, 251, 255] 中有了长足地发展, 推广到了非光滑系统和微分包含.

将 Ekeland 变分原理应用于偏微分方程控制的最优控制问题的最早工作可能是 Li 与 Yao [786, 787] 和 Plotnikov 与 M. Sumin [1084]. 接下来, 对各种由非线性 (大多是半线性) 椭圆和抛物方程控制的偏微分控制问题, 这种方法用来推导最大值原理类型的必要最优条件, 这可见 Arada 与 Raymond [23, 24], Bonnans 与 Casas [131], Casas [225], Casas, Raymond 与 Zidani [226], Casas 与 Yong [227], Fattorini [429, 430, 431, 432, 433], Fattorini 与 Frankowska [434], Fattorini 与 Murphy [435, 436], Li 与 Yong [788, 789], Raymond [1120], Raymond 与 Zidani [1121], 以及其他研究人员的工作. 如前所述, 状态约束半线性双曲方程的 Neumann 边界控制问题的第一个结果由 Mordukhovich 与 Raymond[944] 得到, 这基于 Ekeland 变分原理, 是该文中分析的基本要素.

需要指出, 对控制问题基于 Ekeland 变分原理的逼近过程, 偏微分的情形比常微分的情形要复杂得多. 特别地, 在通过 Ekeland 变分原理推导偏微分问题的必要最优条件的时候, 有界和无界控制的情形是非常不同的. 其中的主要原因与这样一个事实相关, 即在偏微分系统无界控制框架下应用 Ekeland 变分原理的时候, 所需的完备度量空间的完备性很难构造, 而且不容易保证对应的罚函数的下半连续性. 关于无界控制的最早结果由 Fattorini[431], 并独立地由 Raymond 与 Zidani[1121] 在抛物情形得到. Mordukhovich 与 Raymond[944] 建立的方法重述于 7.2 节, 它扩展了文献 [1121] 中的方法, 将问题推广到了 Neumann 边界条件具有无界控制的状态约束双曲系统. 该方法的具体实现很大程度上是基于 Lasiecka 与 Triggiani[744, 745] 关于双曲 Neumann 边值问题正则性理论以及 7.2.2 节的相关结果.

7.5.12 针形扩散控制扰动

对由半线性波动方程控制的状态约束 Neumann 问题的必要最优条件的推导, 7.2 节发展的扰动技术的一个重要部分是对逼近过程中产生的没有状态约束的逼近问题的变分分析. 与 6.3.2 小节用来证明自由端点常微分控制系统的 Pontryagin 最大值原理的涉及针形控制变分的增量方法相比, 在 7.2.3 给出的这种分析所基于的

技术可以看成它的一种高维 (这里的情形是双曲) 版本.

可以看到, 所论的偏微分情形比其常微分版本要复杂得多. 该分析中用到的针形控制变分在偏微分控制文献中也称为“扩散扰动”, 也成为“钉型/多重钉型”或“补丁”变分; 请比照 Fattorini [432], Li 与 Yong [789], 和 Raymond 与 Zidani [1121]. 这种最优控制的变分/扰动最早由 Li 与 Yao[786, 787] 用来推导偏微分控制问题的必要最优条件, 然后在很多文献中得到了发展, 特别可见前面提到的文献 [432, 789, 1121] 及其所录文献. 注意这种扰动所需性质的证明及其在必要最优条件证明中的具体实施都基于 Lyapunov-Aumann 凸化定理, 既隐含地利用了凸性; 见 Raymond 与 Zidani [1121, 引理 4.2], Mordukhovich 与 Raymond [944, 引理 4.2 与定理 4.1], 以及 7.2.3 小节在引理 7.19 和定理 7.18 的证明中重述的结构.

利用扩散控制扰动, 可以推导出双曲方程的 Neumann 边界控制问题 (包括 7.2 节考虑的逼近和状态约束系统) 的逐点形式的 Pontryagin 最大值原理, 这与第 6 章研究的常微分控制问题是类似的. 但是, 对从无约束问题的必要最优条件而得到有约束问题的结果的极限过程, 常微分和双曲偏微分的情形却截然不同 (见 6.2.1 节和 7.2.4 节); 后者极大地基于 7.2.2 节发展的 Neumann 类型问题的弱解的正则性理论.

7.5.13　双曲系统的 Dirichlet 边界控制

7.3 节致力于状态约束 n 维线性拨动方程的 Dirichlet 边界控制问题. 事实上所得结果可以推广到更一般的双曲线性方程, 其中可由强椭圆算子替代经典的 Laplace 算子. 7.3 节基于 Mordukhovich 与 Raymond 的近期文章 [943], 除此而外作者不知道还有别的在该节所论控制问题上的结果.

众所周知, 与所有类型的偏微分系统的分布和 Neumann 边界控制问题相比, Dirichlet 边界控制问题提供了最少的正则性; 请比照 7.4 节中更多的讨论及其对抛物方程的评注. 与偏微分方程的其他类型相比, 双曲情形的正则性又是最少的.

据作者所知, Dirichlet 双曲边值问题的最强的正则性理论是由 Lasiecka, Lions 与 Triggiani[740] 给出的; 对进一步的材料和应用, 可见随后 Lasiecka 与 Sokolowski 的文章 [741], 和 Lasiecka 与 Triggiani 的书 [746]. 为了这里的变分分析 7.3 节应用了这个理论. 但是, 对比 7.2 节考虑的 Neumann 情形和 7.4 节研究的 Dirichlet 双曲情形, 偏微分方程理论的这个正则性理论并不适用于 Dirichlet 边界控制的情形的类似的变分方法和结果. 事实上, 相关的结果是通过别的办法建立起来的, 7.3 节阐述的就是该方向上的成果.

7.3 重述了前述文章 [943] 建立的方法. 该方法基于把状态约束 Dirichlet 边界控制问题约化成具有算子和几何约束的无穷维规划问题. 在双曲 Dirichlet 边界控制情形, 由于缺少正则性, 所以不能进行扰动/逼近分析, 这由极小化费用函数中的

被积项的额外的完整凸性条件来补偿, 从而保证了在所得的无穷维数学规划问题上可以应用 Lagrange 乘子法的某种恰当版本 (可特别比照 Alibert 与 Raymond [9]), 这里需要 7.3.3 小节建立的伴随 Dirichlet 系统弱解的性质. 所给的假设和问题所具有的正则性也用来建立了所论状态约束 Dirichlet 边界控制问题的最优控制的存在性.

7.5.14 优化与控制中的极小极大问题

具有极小极大费用函数的的优化问题在很多课题中都至关重要, 这包括优化和均衡理论的很多方面, 开环和闭环控制问题的分析和综合、各种 (静态和动态) 博弈理论框架以及其他很多应用. 值得再次指出的是, 作为内蕴地非光滑的极小极大函数及相关联的极小极大问题一直是非光滑变分分析和广义微分技术的发展和应用的主要动力之一. 在极小极大理论及其应用的各种课题上的大量文献中, 这里特别指出 Başar与 Bernhard [87], Chernousko [238], Chikrii [239], Danskin [307], Demyanov 与 Malozemov [319], Freeman 与 Kokotović[474], Krasovskii 与 Subbotin [702], Kryazhimskii 与 Osipov [721], Kurzhanskii [731], Kurzhanskii 与 Vályi[732], Moiseev [885], von Neumann 与 Morgenstern [1000], Rockafellar 与 Wets [1165], Subbotin [1230], Simons [1213], 以及其中所录文献.

5.3.2 小节和 5.5.19 小节已经从多目标优化和广义微分的角度讨论了极小极大的某些课题, 而 7.4 节的主要目的是研究具有 Dirichlet 边界条件和逐点状态约束的抛物系统的极小极大控制问题. 之前的由偏微分方程控制的控制问题的某些极小极大问题的研究可见 Ahmed 与 Xiang [5], Arada [21], Arada, Bergounioux 与 Raymond [22], Lenhart, Protopescu 与 Stojanović[762], Li 与 Yong [789], Mordukhovich [905, 918], Mordukhovich 与 Shvartsman [957], Mordukhovich 与 Zhang [978] 等.

7.5.15 约束抛物系统的极小极大控制

7.4 节中极小极大控制问题的研究动力主要来自于不确定条件下土壤水自动调节的变化规律上的应用; 见 Mordukhovich [898, 905]. 这种系统的动态过程由过滤/扩散理论中线性化了的抛物方程来描述, 其具有作用于 Dirichlet 边界条件上的有界控制, 以及放在抛物方程右边的分布型的不确定扰动. 进一步, 通过在受控运动上加逐点状态约束, 这些实际问题在技术上的主要需求就可以得到满足.

由于在文献 [898, 905] 中所建立的工程控制问题模型中没有关于不确定扰动的概率信息, 极小极大似乎是优化和控制设计的最自然的判据. 由此, 文献 [898, 905] 对一维线性抛物方程的情形解决了一些开环和反馈控制问题并给出了实际的操作, 这是通过使用动力系统和约束的特殊性质来实现的. 具有“硬”约束的高维抛物系统则需要更严格的考察, 它推动了 7.4 节给出的研究. 该节内容主要基于

Mordukhovich 与 Zhang [978].

7.4 节研究的极小极大问题 (P) 涉及一个线性高维抛物方程, 它由具有变系数的强一致椭圆算子描述, 而该算子在 Hilbert 空间上生成一个解析半群. 可测控制作用于 Dirichlet 边界条件上, 而不确定的扰动累加地分布在方程上. 具有 Dirichlet 边界条件的线性抛物方程解的概念是按照适度解的意义来理解的, 见 7.4.1 小节和下面更多的注释. 极小极大最优解的概念是标准意义下的, 即所给依赖于控制、扰动和状态变量的积分泛函的鞍点 (最差扰动和最佳控制). 这类极小极大问题的一个显著特征是控制、扰动和轨道上具有的幅度/大小类型的“硬”/逐点约束.

由于该动力系统和所给约束的结构具有线性性, 这就可以把这样的极小极大问题分拆成关于最差扰动和 Dirichlet 边界控制的两个相关联的问题, 并分别用不同的方法来研究. 但是, 两种方法都涉及一定的光滑逼近过程, 它严重依赖于具有 Dirichlet 类型的非正则/可测边界条件的抛物系统适度解的适定性.

7.5.16　具有 Dirichlet 边界条件的抛物系统的适度解及其性质

对具有非正则 (仅可测)Dirichlet 边界条件的抛物系统的适度解, 其研究特别可见 Balakrishnan [74], Lasiecka [739], Lasiecka 与 Triggiani[742, 743], 和 Washburn [1324], 读者可以在这些文献中找到存在性和唯一性的结果以及基本的算子估计 (7.71). 7.4.2 小节的结果主要取材于 Mordukhovich 与 Zhang [978], 并由之建立了定理 7.36 的极小极大解的存在性 (基于该节中使用的恰当拓扑下的经典的 von Neumann 极小极大定理), 进而证明了逼近过程的收敛性/适定性, 该过程在 7.4.3~7.4.5 小节被用来推导必要最优和次最优条件.

这里要注意在极小极大问题 (P) 及其分拆的控制/扰动版本中的一个特别情况, 即状态约束由逐点形式 (7.62) 给出, 只涉及一般的二元连续实值函数 $y(x,t)$, 而所用的半群方法所处理的定义 6.26 中依赖于时间的连续适度解却取值于泛函空间. 定理 7.35 在消除这个差异中起着关键的作用, 它建立了二元状态解按值的几乎处处逐点收敛性, 这是由 Dirichlet 边界条件中的可测函数对应的弱收敛导出的.

7.5.17　具有非正则/非光滑数据的约束抛物系统的分布控制

根据分拆过程, 原极小极大问题的最差扰动恰巧是具有固定 Dirichlet 边界条件和逐点状态约束的抛物分布控制问题的最优解. 这种控制问题的必要最优条件举例可见 Casas [225], Bergounioux 与 Tröltzsch[104], Mackenroth [825], Fattorini [430, 432], Raymond 与 Zidani [1121], Tröltzsch[1271]; 也可参阅其中的文献. 对抛物系统可注意到, 分布控制问题在方法和结果上都类似其具有 Neumann (非 Dirichlet) 边界条件的边界控制版本.

但是, 由于其数据上显著的非正则性, 7.4.3 小节和 7.4.5 小节研究的分布控制

问题 (P_1) 与其标准版本是有本质区别的. 问题出在 (P_1) 中的费用函数和状态约束上, 它们随原极小极大问题的分拆而出现, 但依赖于所给的 Dirichlet 边界控制问题 (P_2) 的适度解 $\bar{y}_2(x,t)$, 其作为两个变量的函数通常具有很强的非连续性. 在 PDE 的文献中这类问题一般称为非光滑的.

为了处理这种具有非正则/非光滑数据的分布控制问题, 这里是根据 Mordukhovich 与 Zhang 的文章 [978], 该文主要基于抛物变分不等式理论中的一个特定的光滑逼近技术; 文献举例可见 Barbu [82], Friedman [479], He [554], Neittaaanmäki与 Tiba [997], Tiba [1255], 及其所录的文献. 7.4.3 小节用的技术与基于 Ekeland 变分原理的方法是不同的, 但同样能对这种非正则状态约束问题有效地使用罚函数而使之变为一个光滑无约束问题的参数族, 并建立所需的近似解的强收敛性, 进而在适定逼近问题中导出必要最优条件.

在 7.4.5 小节中完成的下一个目标是在状态约束分布控制问题 (P_1) 的最优解 (最差扰动) 必要条件的证明中阐明这样一个极限过程, 即对无状态约束参数逼近问题 $(P_{1\varepsilon})$ 的必要最优条件取极限. 在此极限过程的具体操作中, 用到了一个改进的规范条件, 其所涉及的情形与 Bergounioux 与 Tiba [103], Bergounioux 与 Tröltzsch[104] 等文章中的不太相同. 这里的条件 (CQ1) 不需要 (P_1) 中的可行解集具有非空内部, 这与凸规划中经典的松弛约束规范条件在无穷维中的版本是有很大不同的. 基于 (CQ1) 并利用 Brézis 与 Strauss [177] 中的一个精致的压缩结果和经典的 Sobolev 嵌入定理, 由此得到以 Pontryagin 最大值原理的积分形式给出的最差扰动的必要最优条件, 容易导出逐点的砰–砰关系.

7.5.18 具有逐点状态约束的抛物系统的 Dirichlet 边界控制

对极小极大分拆过程中的第二个问题 (P_2) 是一个具有状态约束的 Dirichlet 边界控制问题, 它拥有作用于 Dirichlet 边界条件上的极不正则 (L^∞) 的控制函数. 众所周知, 对抛物方程而言, Dirichlet 边界控制情形是最优控制中最具挑战性的问题, 这由于该条件对抛物动力系统所提供的极少的正则性质. 关于具有 Dirichlet 边界控制的抛物系统的各种必要最优条件, 这里建议读者参阅 Arada 与 Raymond [23, 24], Barbu [82], Fattorini [433], Fattorini 与 Murphy [435, 436], Lasiecka [739], Lasiecka 与 Triggiani [742, 743, 746], Mordukhovich 与 Zhang [978, 979], 以及 Washburn [1324].

7.4.4 小节和 7.4.5 小节关于 Dirichlet 抛物方程问题的研究主要取材于 Mordukhovich 与 Zhang [978, 979]. 这基于 7.4.2 小节给出的适度解的正则/稳定性质, 以及抛物变分不等式理论中的逼近方法, 该方法的精髓类似 7.4.3 小节和 7.4.5 小节在分布控制/扰动性情的相关论述.

注意到, 该方向上以前的结果 (比如见 Barbu 的书 [82] 以及其中的文献) 对 Dirichlet 边界控制都有强得多的光滑性的要求, 至少要假设

$$u \in W_p^{2-1/p,\,1-2/p}(\Sigma), \quad \text{其中} p \geqslant 2.$$

这是因为, 与 7.4 节中的适度解理论不同, 上面提到的以前该方向的工作大多基于抛物 Dirichlet 边界值问题强解的经典理论, 该理论是特别用来处理 Dirichlet 抛物系统 (7.61) 的 $W_p^{2,1}(Q)$ 类型的解 y 的; 见 Ladyzhenskaya, Solonnikov 与 Uralzeva 的书 [735], 这被 Barbu[82] 和其他研究人员大量使用.

在极度非光滑/非正则的情形应用适度解的优点首先是可以像定理 7.36 那样建立极小极大解的存在性, 而且也可对 Mordukhovich 与 Zhang [979] 中研究的更一般的 Dirichlet 边界控制 (非极小极大) 问题的最优控制建立存在性. 注意该结果中去掉了对被积项 $g(x, t, \cdot)$ 的线性假设, 这对极小极大的存在定理是必须的; 见注解 7.37. 进一步, 7.4.2 小节给出的适度解的性质使得在 7.4.4 中可以完成光滑逼近/惩罚过程, 并建立近似最优解的强收敛性, 这不需要十分苛刻的如 Barbu[82] 的光滑假设.

最后, 7.4.5 小节阐明了对具有状态约束的 Dirichlet 边界控制问题 (P_2) 推导出必要最优条件的极限过程; 涉及此类型更一般的最优控制问题的类似结果也可见 Mordukhovich 与 Zhang [979]. 该节用到了约束规范条件 (CQ2), 其精髓类似 Bergounioux 与 Tiba [103], Bergounioux 与 Tröltzsch[104] 中对应的约束规范条件, 只是场合不同; 由此得到的主要结果, 即定理 7.52, 给出了所论的 Dirichlet 边界控制问题的 Pontryagin 最大值原理的一个积分类型的变体, 这是通过 (7.72) 中定义的适度解算子 $\mathcal{L}$ 的伴随算子

$$\mathcal{L}^*\colon L^\infty(Q)^* \to L^\infty(\Sigma)^*$$

来实现的, 其中对偶空间 $L^\infty(Q)^*$ 等同于定义域 Q 上的有界可加函数 (测度) 的集合. 处理作用于 Dirichlet 边界条件上的 L^∞-控制的有点不同的方法和结果见于 Arada 与 Raymond [23, 24], Fattorini [433], 以及 Fattorini 与 Murphy [435, 436].

7.5.19 控制系统的反馈综合/整合和极小极大设计

众所周知, 反馈控制问题, 即当控制函数依赖于状态变量时, 是控制理论中最难, 应用中最重要的问题之一. 在反馈控制设计上有各种方法, 主要是针对由常微分方程控制的系统. 这里不讨论更多的细节, 只指出下述一些近期文献供读者参阅, 这些发展用到了现代变分分析和广义微分的构造和技术: Bardi 与 Capuzzo Dolcetta [85], Cannarsa 与 Sinestrari [217], Clarke, Ledyaev, Sontag 与 Subbotin [263], Clarke, Ledyaev, Stern 与 Wolenski [264, 265], Clarke 与 Stern [269], Fleming 与 Soner [458], Frankowska [472], Freeman 与 Kokotović[474], Goebel [511], Rockafellar 与 Wolenski [1166, 1167], Subbotin [1230], Sontag [1220], Zelikin 与 melnikov [1359], 以及其中所

录文献. 但是, 这些文献中的绝大多数结果都是理论性的, 在多少实际一点的问题中的具体实现则总是需要特殊的考虑和额外的考察.

在实际应用中出现的许多情形中, 控制系统运作于不确定的条件下, 其中没有不确定的干扰/扰动的 (确定的或有随机效应的) 信息, 只知道其可能的变化范围. 极小极大方法, 或是“确保结果原理”, 为在这样的不确定条件下反馈控制系统的设计和综合铺平了自然的道路. 现在普遍认为对策论方法为极小极大控制设计提供了一个一般的框架; 此方向的文献举例可见 Başar与 Bernhard [87], Chernousko [238], Chikrii [239], Krasovskii 与 Subbotin [702], Kryazhimskii 与 Osipov [721], Kurzhanskii [731], Kurzhanskii 与 Vályi[732], Subbotin [1230], 以及其中的文献. 可以看到, 在把对策论的方法和结果应用于具体的控制系统的过程中, 总是需要额外考虑所论问题的特殊性质.

如 7.4 节开头提到的 (后续讨论见 7.5.14), 这里的关于极小极大控制问题研究的动力源于土壤水系统调控中涉及自动恢复的工程设计这个实际问题. 见 Mordukhovich [898, 905]. 该文章中对该问题找到并描述了足够的数学模型, 即状态约束抛物系统的极小极大综合的问题, 它具有分布的不确定扰动和作用于 Dirichlet 边界条件上的控制. 一个典型的该类问题可见注释 7.54. 由于该问题的一般理论上以前没有任何的方法和结果, 这里在动态由一维热方程描述的情形对其解发展了特殊的技术. 这些技术涉及数个逼近过程; 除必要和充分最优条件外, 还特别地利用了一维抛物动态的一些单调性质和无穷水平上的渐近性质. 另外进一步还知道这些系统表现出一定的大道性质 (turnpike behavior) (如 Carlson, Haurie 与 Leizarowitz [224], Dyukalov [379], Zaslavski [1357] 等书中所描述), 这对文献 [905] 中有效的反馈控制设计是至关重要的.

文献 [898, 905] 中创始的一些方法和结果的一些高维拓展见 Mordukhovich [918], 和 Mordukhovich 与 Shvartsman [957]. 但是, 所论的约束抛物系统的极小极大设计现在还几乎没什么进展.

第 8 章　经济学应用

本书的最后一章致力于现代变分分析和广义微分技术在福利经济学的竞争均衡模型上的应用, 这类模型涉及具有无限维商品空间的非凸经济. 注意, 对于最优化理论、变分方法和广义微分结构的应用来说, 经济模型一直就是一个具有挑战性的领域. 特别地, 本章所考虑的这种类型的福利经济的凸模型, 是 20 世纪 50 年代初凸分析发展的最重要的动因. 从那时起, 这些模型就一直是高等变分和广义可微技术在凸和非凸情形应用上的一个诱人领域.

在研究福利经济的非凸模型中, 主要工具是变分分析的极点原理, 它使得可以建立关于具有边际/均衡价格的非凸经济体系中的弱 Pareto, Pareto, 强 Pareto 最优配置的新的所谓的"广义/扩展的第二福利定理". 这里的边际/均衡价格通过本书探讨的基本法锥及其类 Fréchet 近似来形式化.

8.1　福利经济学模型

这一节在经典和高等的框架下描述福利经济模型, 定义相应的均衡和 Pareto 类型最优概念, 并讨论随后研究的 Pareto 和弱 Pareto(非强 Pareto) 最优配置所需的所谓"净需求规范条件". 这里以一种非正式 (然后正式) 的方式开始描述这些模型及其研究和应用中用到的数学技巧.

8.1.1　基本概念和模型描述

福利经济的经典 Walras 均衡模型及其各种推广长期以来一直被公认为是经济理论和应用中的一个重要部分. Pareto 效率/Pareto 最优概念及其变体对于均衡研究以及对竞争经济作出最优决断起着至关重要的作用.

研究具有光滑数据的经济模型的 Pareto 最优的一个经典方法是将其简约为传统的数学规划问题, 并利用涉及 Lagrange 乘子的一阶必要最优条件. 以这种方式, 在 20 世纪 30 年代末和 40 年代里, 得到了一些重要的结论, 那时已经被证实在任何资源 Pareto 最优配置中消费和生产的边际替代率是彼此相等的; 参见基础性著作 Samuelson[1188] 中的讨论和参考文献, 以及本章末的进一步的评注.

20 世纪 50 年代初, Arrow[26] 和 Debreu[309] 考虑可具有非光滑但凸的数据的经济模型从而把福利经济学理论推进了至关重要的一步. 基于经典的凸集分离定理, 他们及其追随者给出了一个漂亮的理论, 这个理论特别地包含了 Pareto 最优

配置的充分必要条件, 并指出每一个这样的配置都导出凸经济的一个分散均衡. 这个理论的关键结论就是众所周知的经典的福利经济的第二基本定理, 它指出, 任何 Pareto 最优配置都是分散价格均衡, 即, 它可以由一个非零价格向量承载, 此时每个消费者可以最大限度减少支出, 每个企业获得最大利润. 这一结果的完整版本无疑基于凸性, 这对于 Arrow-Debreu 模型及其基于凸分析的扩展至关重要. 还要注意, 福利经济学中 Arrow-Debreu 一般均衡理论, 对作为一门数学学科的凸分析及其后续的各种应用的发展发挥了重要的刺激作用.

另一方面, 凸性相关假设在许多重要的应用中是否有意义往往是有疑问的, 这在 Arrow-Debreu 模型形成之前就被人们认识到了; 例如, 上面提到的 Samuelson 的书 [1188, 231 – 232页] 指出, 这种假设能实现时"纯属巧合 ···". 众所周知, 特别当生产部门规模收益递增时, 凸性假设是站不住脚的. 研究非凸模型的一个常用方法是基于局部切向凸逼近然后采用经典的凸锥分离定理. 这已经通过利用 Clarke 切锥这一结构得以实现, 该结构自动就是凸的. 以这种方式, 边际价格形式化为 Clarke 法锥的对偶, 然而, 该对偶有时会太大而不能在非凸模型中得出满意的结果, 且常常不能对边际成本定价给出任何限制. 读者可以在 Khan 的论文 [671] 中找到许多例子、讨论和参考文献. 该文包含了在有限维商品空间非凸经济第二福利定理更充分的扩展, 其中边际价格通过我们的 (非凸) 基本法锥来表示. 其方法是首先在适当的约束规范条件下, 把 Pareto 最优配置约化成一个不可微规划问题的最优解, 然后应用 Mordukhovich[892] 建立的非光滑优化的必要条件. 这种方法不需要使用凸分离和/或相关的凸分析结果.

本章的基本目标是在福利经济学的非凸模型中建立有关第二福利定理的扩展的全面结果. 这里的基本框架是无穷维商品空间, 所有的结果都基于极点原理, 它是本书中变分分析的主要工具. 正如第 2 章所探讨, 极点原理可以被视为一个在非凸情形中 (局部) 分离的变分变体. 另一方面, 它为非凸集极值点提供了必要的条件. 如后所述, 它包括了 Pareto 最优配置. 因此, 运用极点原理, 这里其实统一了上面讨论的两种方法, 这些方法或是基于将 Pareto 最优归结为数学规划或应用凸集分离定理.

由第 2 章发展的极点原理可获得近似/模糊和确切/极限两种形式的非凸经济第二福利定理扩展版本, 这只需要净需求规范条件这样不太强的假设, 在 Pareto 和弱 Pareto 的最优配置的情形是必须的. 由此便得到了当商品空间有序的情况下确保边际价格正性的有效条件. 所获得的结果即使在凸经济中也带来新的信息, 因为我们不要求任何经典的内部条件或广泛实施的由 Mas-Colell[855] 提出的适宜条件. 此外, 与具有有序商品空间的凸经济的大多数出版物不同, 这里的方法不需要任何有限维或无限维商品空间的格结构.

对具有有序商品空间的非凸经济模型中的强 Pareto 最优配置, 由极点原理在

近似和确切两种形式下都可获得广义第二福利定理的一些出人意料的结果. 事实上, 在这种情况下, 不需要上述规范条件第二福利定理就成立. 这个结论甚至对涉及有限维商品经济的凸经典模型似乎也是新的.

如上所述, 第二福利定理非凸扩展中边际价格的确切版本由基本法锥表示, 近似版本由类 Fréchet 结构表示, 见下面所有三种 Pareto 最优配置的表述. 因此, 由 1.1.4 节建立的近似法锥的变分描述, 可为涉及非线性价格非凸经济学中广义第二福利定理, 得出一个凸型/分散均衡解释; 参看 8.2 节更多细节和讨论.

接下来正式描述本章探讨的福利经济学的基本模型. 虽然给出的描述和本节随后讨论的性质在具有局部凸 Hausdorff 拓扑的线性拓扑空间的一般框架中成立, 但是涉及广义法向量的主要结果需要 Asplund 空间结构; 在其他类型的 Banach 空间上, 这些结果的一些变体亦可见 8.4 节.

设 E 为经济 $\mathcal{E}$ 的赋范商品空间, 它涉及具有消费集合 $C_i \subset E, i = 1, \cdots, n$ 的 $n \in \mathbb{N}$ 个消费者, 以及具有产品集合 $S_j \subset E, j = 1, \cdots, m$ 的 $m \in \mathbb{N}$ 个厂商/企业. 每一个消费者有偏好集 $P_i(x)$, 它包括消费者在消费计划/束 $x = (x_1, \cdots, x_n) \in C_1 \times \cdots \times C_n$ 对 x_i 有偏好的 C_i 中的元素. 对由如 5.3 节中的偏好 $\prec_i$, 特别是福利经济学经典模型中效用函数给出的序关系, 偏好集是一个有价值的推广 (拥有有意义的经济解释). 按照定义可以得出, 对于 $i = 1, \cdots, n$ 有 $x_i \notin P_i(x)$, 另外总假设对于某些 $i \in \{1, \cdots, n\}$ 成立 $P_i(x) \neq \varnothing$, 即, 至少一个消费者是不饱足的. 为方便起见, 若 $P_i(x) = \varnothing$, 置 $\mathrm{cl} P_i(x) := \{x_i\}$.

现在通过商品空间的一个给定的非空子集 $W \subset E$, 给实行市场约束的经济$\mathcal{E}$的可行配置下个定义; 在 $\mathcal{E}$ 中把净需求约束集记为 W.

定义 8.1(可行配置) 设 $x = (x_i) := (x_1, \cdots, x_n)$, $y = (y_j) := (y_1, \cdots, y_m)$. 称二元组 $(x, y) \in \prod_{i=1}^n C_i \times \prod_{j=1}^m S_j$ 为 $\mathcal{E}$ 的一个"可行配置", 如果

$$w := \sum_{i=1}^n x_i - \sum_{j=1}^m y_j \in W. \tag{8.1}$$

引入净需求约束集可以统一一些经济模型的传统情形, 并在一般框架下提供了有益的经济上的意义. 事实上, 在经典情形中, 集合 W 只包含一个元素 $\{\omega\}$, 其中 ω 是一个稀缺资源的总禀赋. 此时约束 (8.1) 归结为"市场出清" 的条件. (8.1) 也包括另一个传统框架, 即当商品空间 E 的序由闭正锥 E_+ 给出时, 置 $W := \omega - E_+$, 这对应于商品的"隐含自由处置". 一般来讲, 约束 (8.1) 描述的是一些很自然的情形, 这特别包括初始总禀赋由于种种原因 (比如信息不全) 而不明确的情况. 此时集合 W 一般地是反映了所考虑的经济模型的某些不确定性.

接下来主要关注经济模型 $\mathcal{E}$ 中可行配置最优化的三个 Pareto 类型概念: 弱

Pareto 最优、Pareto 最优和强 Pareto 最优. 前两个概念在类似但略有不同的净需求约束条件中, 将被视为平行的, 强 Pareto 最优在有序商品空间的情况下起到特定的作用, 这种约束条件, 对于下面建立的第二福利定理的扩展版本, 尤其对于带有限维商品的凸经济的经典框架来讲, 都是不必要的.

定义 8.2(Pareto 型最优配置) 设 $(\bar{x},\bar{y})$ 为经济 $\mathcal{E}$ 中具有以下特点的可行配置:

$$\bar{x}_i \in \text{cl}P_i(\bar{x}), \quad i=1,\cdots,n,$$

我们说:

(i) 如果存在 $(\bar{x},\bar{y})$ 的一个邻域 O, 使得对于每个可行配置 $(x,y)\in O$, 存在某 $i\in\{1,\cdots,n\}$, 满足 $x_i\notin P_i(\bar{x})$, 则称 $(\bar{x},\bar{y})$ 是经济 $\mathcal{E}$的一个局部弱 Pareto 最优配置.

(ii) 如果存在 $(\bar{x},\bar{y})$ 的一个邻域 O, 使得对于每个可行配置 $(x,y)\in O$, 或对于某 $i\in\{1,\cdots,n\}$ 有 $x_i\notin \text{cl}P_i(\bar{x})$; 或对于所有 $i=1,\cdots,n$, 有 $x_i\notin P_i(\bar{x})$, 则称 $(\bar{x},\bar{y})$ 是经济 $\mathcal{E}$ 的一个局部 Pareto 最优配置.

(iii) 如果存在 $(\bar{x},\bar{y})$ 的一个邻域 O, 使得对于满足 $(x,y)\neq(\bar{x},\bar{y})$ 的所有可行配置 $(x,y)\in O$, 对于某 $i\in\{1,\cdots,n\}$, 有 $x_i\notin \text{cl}P_i(\bar{x})$, 则称 $(\bar{x},\bar{y})$ 是经济 $\mathcal{E}$ 的一个局部强 Pareto 最优配置.

当偏好集 $P_i(x)$ 通过偏好关系 $\prec_i$ 定义, 如 5.3 节 (尤其是通过效用函数), 则上述 Pareto 和弱 Pareto 最优配置概念归结为特殊类型约束 (8.1) 下的多目标优化相应的概念. 强 Pareto 最优概念在多目标优化中是非传统的, 即使在经典的框架下亦如此, 然而在经济建模中却起着重要的作用.

为研究 Pareto 和弱 Pareto 最优配置, 在下一小节引入和讨论了恰当的净需求规范条件, 由此能够把该类 Pareto 最优问题归结为某些闭集的局部极值点. 在强 Pareto 最优配置中这种规范条件并不需要, 这一点将在 8.3.2 小节说明.

8.1.2 Pareto 和弱 Pareto 最优配置净需求规范条件

本节从经济$\mathcal{E}$的两个平行的规范条件概念开始, 它们分别在下述的 Pareto 和弱 Pareto 最优配置条件下第二福利定理的扩展中发挥关键作用. 很明显, (ii) 中的条件蕴涵 (i) 中的条件, 但反之不然.

定义 8.3(净需求规范条件) 设 $(\bar{x},\bar{y})$ 为经济 $\mathcal{E}$ 中的一个可行配置, 且置

$$\bar{w} := \sum_{i=1}^{n}\bar{x}_i - \sum_{j=1}^{m}\bar{y}_j \ . \tag{8.2}$$

给定 $\varepsilon>0$, 考虑集合

$$\Delta_\varepsilon := \sum_{i=1}^{n}\text{cl}P_i(\bar{x})\cap(\bar{x}_i+\varepsilon\mathbb{B}) - \sum_{j=1}^{m}\text{cl}S_j\cap(\bar{y}_j+\varepsilon\mathbb{B}) - \text{cl}W\cap(\bar{w}+\varepsilon\mathbb{B}),$$

则说:

(i) "净需求规范 (NDQ) 条件" 在 $(\bar{x},\bar{y})$ 成立, 如果存在 $\varepsilon>0$, 序列 $\{e_k\}\subset E$, 当 $k\to\infty$ 时, $e_k\to 0$, 且有消费者下标 $i_0\in\{1,\cdots,n\}$, 使得对于所有足够大的 $k\in\mathbb{N}$, 有

$$\Delta_\varepsilon+e_k\subset P_{i_0}(\bar{x})+\sum_{i\neq i_0}\mathrm{cl}P_i(\bar{x})-\sum_{j=1}^m S_j-W. \tag{8.3}$$

(ii) "净需求弱规范 (NDWQ) 条件" 在 $(\bar{x},\bar{y})$ 成立, 如果有 $\varepsilon>0$, 当 $k\to\infty$ 时, 序列 $e_k\to 0$, 使得对于所有足够大的 $k\in\mathbb{N}$, 有

$$\Delta_\varepsilon+e_k\subset\sum_{i=1}^n P_i(\bar{x})-\sum_{j=1}^m S_j-W. \tag{8.4}$$

容易看到, 如果在定义 1.24(ii) 意义上偏好、生产或净需求约束集中的一个在相应点周围是上图 Lipschitz 的, 则 NDQ 和 NDWQ 两个条件自动成立. 从命题 1.25 知道, 凸集 $\Omega\subset X$ 的上图 Lipschitz 性质等价于其内部非空, 即 $\mathrm{int}\Omega\neq\varnothing$. 这样, 上述规范条件可以被视为福利经济学凸模型中发展的经典非空内部条件的深远扩展.

下一个命题中包含了能确保实现 NDQ 和 NDWQ 性质的可以验证的条件, 这些条件极大地推广了上述的上图 Lipschitz 假设. 注意此时 Ω 在 $\bar{x}$ 周围的上图 Lipschitz 性质蕴涵其闭包 $\mathrm{cl}\Omega$ 在该点附近的相同性质, 但反之不然. 另外值得一提的是, 如在下面 (8.6) 和 (8.7) 中的情形, 集合的求和 (尤其是大量的集合) 趋于改善与非空内部相关的性质, 集合的上图 Lipschitz 性质就属于这一类.

定理 8.4(NDQ 和 NDWQ 性质的充分条件)　设 $(\bar{x},\bar{y})$ 为经济 $\mathcal{E}$ 的可行配置. 下列断言成立:

(i) 假设集合 $S_j, j=1,\cdots,m$, W 分别在点 $\bar{y}_j$ 和 (8.2) 中给出的 $\bar{w}$ 附近是闭的. 那么, 如果存在数 $\varepsilon>0$ 、下标 $i\in\{1,\cdots,n\}$ 以及满足当 $k\to\infty$ 时, $e_{ik}\to 0$ 的期望序列 $\{e_{ik}\}\subset E$, 使得对于所有大的 $k\in\mathbb{N}$, 都有

$$\mathrm{cl}P_i(\bar{x})\cap(\bar{x}_i+\varepsilon\mathbb{B})+e_{ik}\subset P_i(\bar{x}), \tag{8.5}$$

则 NDQ 条件在 $(\bar{x},\bar{y})$ 处成立.

此外, 如果在 (8.5) 中, 对于每个 $i\in\{1,\cdots,n\}$, 期望序列 $\{e_{ik}\}$ 对某 $\varepsilon>0$ 存在, 则 NDQ 条件在 $(\bar{x},\bar{y})$ 是成立的.

(ii) 假设对于所有的 $i=1,\cdots,n$, $\bar{x}_i\in\mathrm{cl}P_i(\bar{x})$. 那么 NDWQ 条件在 $(\bar{x},\bar{y})$ 是满足的, 如果集合

$$\Delta:=\sum_{i=1}^n P_i(\bar{x})-\sum_{j=1}^m S_j-W \tag{8.6}$$

在 $0 \in \mathrm{cl}\Delta$ 附近是上图 Lipschitz 的. 这特别包括集合 $P_i(\bar{x})$ $(i = 1, \cdots, n)$, S_j $(j = 1, \cdots, m)$, 以及 W 中的一个, 或者 (8.6) 中组合的一部分在相应点是上图 Lipschitz 的情形.

(iii) 假设 $n > 1$. 则 NDQ 条件在 $(\bar{x}, \bar{y})$ 是满足的, 如果存在消费者下标 $i_0 \in \{1, \cdots, n\}$, 满足 $P_{i_0}(\bar{x}) \neq \varnothing$, 并且集合

$$\Sigma := \sum_{i \neq i_0} \mathrm{cl} P_i(\bar{x}) \tag{8.7}$$

在 $\sum\limits_{i \neq i_0} \bar{x}_i$ 附近是上图 Lipschitz 的. 这特别包括集合 $\mathrm{cl}P_i(\bar{x})$ $(i \in \{1, \cdots, n\} \backslash \{i_0\})$ 中的一个, 或这其中某些集合的和在相应点附近是上图 Lipschitz 的情形.

证明 由定义和假设, (i) 中的两个断言都很容易得出.

现证明 (ii). 由 (8.4) 的构造, 只需考虑当 (8.6) 中的聚合集 Δ 在原点附近是上图 Lipschitz 的情形就够了. 利用上图 Lipschitz 性质的定义 1.24(ii), 可找到 $v \in E$ 和 $\gamma > 0$ 满足

$$\Delta \cap (\gamma \mathbb{B}) + t(v + \gamma \mathbb{B}) \subset \Delta, \quad \forall t \in (0, \gamma). \tag{8.8}$$

选择任意一个数列 $t_k \downarrow 0$ (当 $k \to \infty$ 时), 置

$$e_k := t_k v \quad (k \in \mathbb{N}), \quad \varepsilon := \frac{\gamma}{n + m + 2}. \tag{8.9}$$

下面证明 NDWQ 条件 (8.6) 对 (8.9) 中的 e_k 和 ε 成立. 为此任取 $z_\varepsilon \in \Delta_\varepsilon$, 则由定义 8.3 中 $\bar{w}$ 和 Δ_ε 结构可得 $z_\varepsilon \in (n + m + 1)\varepsilon \mathbb{B}$. 由 (8.6) 中 Δ 的结构, 可找到 $k \to \infty$ 时收敛于 z_ε 的序列 $z_k \in \Delta$. 很明显, 根据 (8.9) 中的 ε 的选择, 对于足够大的 $k \in \mathbb{N}$ 有

$$z_k \in (n + m + 2)\varepsilon \mathbb{B} = \gamma \mathbb{B}. \tag{8.10}$$

另外还可选择 z_k 使得对于足够大的 $k \in \mathbb{N}$ 有

$$z_\varepsilon - z_k \in (t_k \gamma) \mathbb{B}. \tag{8.11}$$

现在将 (8.8)~(8.11) 结合起来, 即得

$$z_\varepsilon + e_k = z_k + t_k v + (z_\varepsilon - z_k) \in \Delta \cap (\gamma \mathbb{B}) + t_k(v + \gamma \mathbb{B}) \subset \Delta,$$

这当然包涵 (8.4).

最后需在给定条件下证明 (iii). 由 (8.7) 中 Σ 的上图 Lipschitz 性质, 可找到 $v \in E, \gamma > 0$ 满足

$$\sum_{i \neq i_0} \mathrm{cl} P_i(\bar{x}) \bigcap \Big(\sum_{i \neq i_0} \bar{x}_i + \gamma \mathbb{B} \Big) + t(v + \gamma \mathbb{B}) \subset \sum_{i \neq i_0} \mathrm{cl} P_i(\bar{x}). \tag{8.12}$$

现在如 (8.9) 选择 v_k 和 ε 并进行如上述 (ii) 中类似的论证. 取 $z_\varepsilon \in \Delta_\varepsilon$ 满足

$$z_\varepsilon = \sum_{i=1}^{n} x_i - \sum_{j=1}^{m} y_j - w, \quad x_i \in \mathrm{cl}P_i(\bar{x}), \quad y_j \in \mathrm{cl}S_j, \quad w \in \mathrm{cl}W,$$

并由相应集合 $P_{i_0}(\bar{x})$, S_j, W 中的元素序列逼近 x_{i_0}, y_j 和 w. 这里与证明 (ii) 不同的是, 对 $i \neq i_0$, 不用逼近 x_i. 由此通过类似论证 (ii) 的方式即可从上图 Lipschitz 性质 (8.8) 导出净需求规范条件 (8.3). 这给出了 (iii) 并完成命题的证明. △

需要看到的很重要的一点是, 如果净需求约束集 W 在 $\bar{w}$ 附近是上图 Lipschitz 的, 那么为了实现定义 8.3 的两个规范条件, 并不需要给偏好和生产集加任何条件. 这很容易从命题 8.4(ii) 得出. 这特别包括当 E 是有序的且对于闭正锥 $E_+ \subset E$, 有 $W := \omega - E_+$ (其中 $E_+ \neq \varnothing$) 的情形. 该情形涵盖了 Cornet[288] 定义的所谓“自由处置 Pareto 最优化”这个传统情形.

8.2　非凸经济学的第二福利定理

本节包含非凸型经济 $\mathcal{E}$ 的 Pareto 和弱 Pareto 最优配置的必要条件, 所涉及的商品空间 E 是 Asplund 空间, 但不假设任何序结构. 利用第 2 章的极点原理, 以近似和确切的形式, 分别通过预法锥 /Fréchet 法锥和基本法锥, 获得了两种 Pareto 最优配置的必要条件. 所获得的结果把广义第二福利定理适当地扩展到了对于所有偏好和生产集涉及相同的 (共同) 边缘/均衡价格的非凸经济. 本节还讨论了这些主要结果的各种后继结果和解释, 其中包括一些相当令人惊讶的结果, 这些结果通过采用非线性价格, 确保了非凸模型的凸类型的分散均衡.

8.2.1　第二福利定理的近似版本

这一小节探讨扩展第二福利定理的近似/模糊版本, 这在 Pareto 和弱 Pareto 最优配置的表述和证明上是并行给出的.

定理 8.5 (具有 Asplund 商品空间的扩展第二福利定理的近似形式)　设二元组 $(\bar{x}, \bar{y})$ 为具有 Asplund 商品空间的经济 $\mathcal{E}$ 的局部 Pareto (对应地, 弱 Pareto) 最优配置. 假设净需求规范条件 (对应地, 净需求弱规范条件) 在 $(\bar{x}, \bar{y})$ 得到满足. 那么对于每一个 $\varepsilon > 0$, 都存在次最优三元组:

$$(x, y, w) \in \prod_{i=1}^{n} \mathrm{cl}P_i(\bar{x}) \times \prod_{j=1}^{m} \mathrm{cl}S_j \times \mathrm{cl}W$$

(其中 w 定义于 (8.1)) 和共同的边际价格 $p^* \in E^* \setminus \{0\}$ 满足如下的 (8.13)~(8.16):

$$-p^* \in \widehat{N}\big(x_i; \mathrm{cl}P_i(\bar{x})\big) + \varepsilon \mathbb{B}^*, \tag{8.13}$$

这里, 对所有的 $i=1,\cdots,n$, 有 $x_i\in\bar{x}_i+\frac{\varepsilon}{2}\mathbb{B}$,

$$p^*\in\widehat{N}(y_j;\mathrm{cl}S_j)+\varepsilon\mathbb{B}^*,\tag{8.14}$$

这里, 对所有的 $j=1,\cdots,m$, 有 $y_j\in\bar{y}_j+\frac{\varepsilon}{2}\mathbb{B}$,

$$p^*\in\widehat{N}(w;\mathrm{cl}W)+\varepsilon\mathbb{B}^*,\tag{8.15}$$

这里, $w\in\bar{w}+\frac{\varepsilon}{2}\mathbb{B}$, 并且

$$\frac{1-\varepsilon}{2\sqrt{n+m+1}}\leqslant\|p^*\|\leqslant\frac{1+\varepsilon}{2\sqrt{n+m+1}},\tag{8.16}$$

其中, $\bar{w}$ 是在 (8.2) 中定义的.

证明 设 $(\bar{x},\bar{y})$ 为经济 $\mathcal{E}$ 的可行配置. 假设这种配置在定义 8.2 中的 Pareto 或弱 Pareto 意义下属于局部最优, 下面就这两种情形并行地给出证明. 事实上, 这两种情形的唯一区别是运用来自定义 8.3 的相应的净需求规范条件, 旨在把所考虑的 Pareto 最优归结为一个特别的集合系统的局部极值点. 考虑具有下面范数的生产空间 $X:=E^{n+m+1}$:

$$\|(v_1,\cdots,v_{n+m+1})\|_X:=\left[\|v_1\|^2+\cdots+\|v_{n+m+1}\|^2\right]^{\frac{1}{2}}.$$

因为 E 为 Asplund 空间, 生产空间 X 亦为 Asplund 空间. 现取 $\varepsilon>0$ 使得 NDQ(对应地, NDWQ) 条件成立, 其中 (8.3) 和 (8.4) 中相应的序列为 $\{e_k\}$. 定义 X 中两个闭集如下:

$$\Omega_1:=\prod_{i=1}^{n}\left[\mathrm{cl}P_i(\bar{x})\cap(\bar{x}_i+\varepsilon\mathbb{B})\right]\times\prod_{j=1}^{m}\left[\mathrm{cl}S_j\cap(\bar{y}_j+\varepsilon\mathbb{B})\right]\times\left[\mathrm{cl}W\cap(\bar{w}+\varepsilon\mathbb{B})\right],\tag{8.17}$$

$$\Omega_2:=\left\{(x,y,w)\in X\,\middle|\,\sum_{i=1}^{n}x_i-\sum_{j=1}^{m}y_j-w=0\right\}.\tag{8.18}$$

可验证 $(\bar{x},\bar{y},\bar{w})$ 是集合系统 $\{\Omega_1,\Omega_2\}$ 的局部极值点. 事实上, 从 (8.1) 和 (8.2) 直接得出 $(\bar{x},\bar{y},\bar{w})\in\Omega_1\cap\Omega_2$. 为证明 $(\bar{x},\bar{y},\bar{w})$ 的局部极值点性质, 只要找到该点的邻域 U 和序列 $\{a_k\}\subset X$ 就足够了, 这里 a_k 满足当 $k\to\infty$ 时, $a_k\to 0$, 并且对于所有大的 $k\in\mathbb{N}$, 有

$$(\Omega_1-a_k)\cap\Omega_2\cap U=\varnothing.\tag{8.19}$$

为此取 Pareto(弱 Pareto) 配置的邻域 $O\subset E^{n+m}$, 一个分别保证 (8.3) 和 (8.4) 的趋于零的序列 $\{e_k\}\subset E$. 在两种情况下都置

$$U:=O\times\mathbb{R}\subset X \text{ 和 } a_k:=(0,\cdots,0,e_k)\in X.$$

下证对于 (8.3) 和 (8.4) 中同样的 $k \in \mathbb{N}$, (8.19) 成立. 否则可找到 $z_k \in \Omega_1$ 满足 $z_k - a_k \in \Omega_2$. 根据 (8.17) 和 (8.18) 的结构以及 a_k 和 U 的构造, 这意味着存在 (x_k, y_k, w_k), 其中 $(x_k, y_k) \in O$, 满足

$$x_{ik} \in \mathrm{cl}P_i(\bar{x}) \cap (\bar{x}_i + \varepsilon\mathbb{B}), \quad i = 1, \cdots, n,$$

$$y_{jk} \in \mathrm{cl}S_j \cap (\bar{y}_j + \varepsilon\mathbb{B}), \quad j = 1, \cdots, m,$$

$$w_k \in \mathrm{cl}W \cap (\bar{w} + \varepsilon\mathbb{B})$$

和

$$\sum_{i=1}^{n} x_{ik} - \sum_{j=1}^{m} y_{jk} - w_k + e_k = 0.$$

根据定义 8.3 中集合 Δ_ε 的构造, 后者蕴涵 $0 \in \Delta_\varepsilon + e_k$. 然后利用 NDQ 条件知原点属于 (8.3) 式右边的集合, 而 NDWQ 条件保证了 (8.4) 中右边集合包含原点. 这无疑与第一种情形中 $(\bar{x}, \bar{y})$ 的 (局部)Pareto 最优性、与第二种情形中 $(\bar{x}, \bar{y})$ 的弱 Pareto 最优性矛盾. 这样就得到了 (8.19), 这表明 $(\bar{x}, \bar{y}, \bar{w})$ 是 Asplund 空间 X 中所考虑的闭集系统 $\{\Omega_1, \Omega_2\}$ 的局部极值点.

现在对该闭集系统运用定理 2.20 中极点原理的近似版本. 则对任意 $\varepsilon > 0$, 都有 $z := (x_1, \cdots, x_n, y_1, \cdots, y_m, w) \in \Omega_1, \widetilde{z} \in \Omega_2$, 和对偶元素 (Fréchet 法向量)

$$z^* \in \widehat{N}(z; \Omega_1) \text{ 与 } \widetilde{z}^* \in \widehat{N}(\widetilde{z}; \Omega_2) \tag{8.20}$$

对于 $i = 1, \cdots, n$, $j = 1, \cdots, m$ 满足条件:

$$\|x_i - \bar{x}_i\| \leqslant \frac{\varepsilon}{2}, \quad \|y_j - \bar{y}_j\| \leqslant \frac{\varepsilon}{2}, \quad \|w - \bar{w}\| \leqslant \frac{\varepsilon}{2}, \tag{8.21}$$

其中

$$\frac{1-\varepsilon}{2} \leqslant \| \widetilde{z}^* \| \leqslant \frac{1+\varepsilon}{2} \quad \text{且} \quad \| z^* + \widetilde{z}^* \| \leqslant \frac{\varepsilon}{2} \quad . \tag{8.22}$$

注意到 (8.18) 中集合 Ω_2 是一个对所有变量 (x_i, y_j, w) 可分离的线性子空间. 由此得 $\widehat{N}(\widetilde{z}; \Omega_2)$ 是与 Ω_2 正交的子空间, 并且在 (8.20) 中

$$\widetilde{z}^* = (p^*, \cdots, p^*, -p^*, \cdots, -p^*),$$

其中带负号的项开始于第 $(n+1)$ 个分量. 由 (8.22) 和 X 上范数的定义得

$$\frac{1-\varepsilon}{2} \leqslant \sqrt{n+m+1} \, \| p^* \| \leqslant \frac{1+\varepsilon}{2} \quad . \tag{8.23}$$

从 (8.18) 和 (8.22) 中最后的那个估计得

$$-\widetilde{z}^* = (-p^*, \cdots, -p^*, p^*, \cdots, p^*) \in \widehat{N}(z; \Omega_1) + \varepsilon\mathbb{B}^*. \tag{8.24}$$

现在把命题 1.2 中的 Fréchet 法锥乘积公式应用到集合 Ω_1 上, 由 (8.21) 观察到 (8.24) 中 z 点的所有分量 (x_i, y_j, w) 都属于 (8.17) 中相应邻域的内部; 因此, 这些邻域在计算 $\widehat{N}(z;\Omega_1)$ 时可以被忽略. 最后结合 (8.21), (8.23) 和 (8.24), 得到关系 (8.13)~(8.16), 并且完成了定理的证明. △

可以看到, 与针对一般闭集极值系统的定理 2.20 中的近似极点原理不同, 定理 8.5 保证, 对于 (8.13)~(8.16) 中涉及的所有集合存在一个共同的对偶元素 $p^* \in E^* \setminus \{0\}$, 而不是极点原理中的给出的一般来说不同的元素 x_i^*. 这一共同的元素可以被解释为在 Pareto 和弱 Pareto 最优配置附近, 所有偏好和生产集的一个近似边际/均衡价格. 这与经典第二福利定理本质对应, 即确保消费者和企业的边际替代率的同一性. 注意, 福利经济学模型中的极点原理的这种特殊情况, 主要是由于极值系统中集合的特殊结构, 特别是由于 (8.18) 中变量的可分离性.

对于凸经济, 即具有凸偏好和生产集合的经济, 下面给出上面建立的第二福利定理近似版本的一个均衡解释. 在这种情况下, 关系 (8.13) 和 (8.14) 分别归结为相应的偏好 (生产) 集合的扰动消费支出 (公司利润) 的全局最小 (最大) 化. 由此而得出小扰动下凸模型的分散价格均衡, 这里并不需要所论问题的凸偏好和生产集的内点假设.

推论 8.6 (凸经济学的扰动均衡) 除了定理 8.5 中的假设之外, 假设所有偏好和产品集合 $P_i(\bar{x})$, $i = 1, \cdots, n$ 和 S_j, $j = 1, \cdots, m$ 是凸的. 那么对于每个 $\varepsilon > 0$, 都存在一个次最优配置 (x, y), 其中可行线性配置来自 (8.1), 满足

$$(x, y, w) \in \prod_{i=1}^{n} \left[\mathrm{cl}P_i(\bar{x}) \bigcap (\bar{x}_i + \frac{\varepsilon}{2}\mathbb{B})\right] \times \prod_{j=1}^{m} \left[\mathrm{cl}S_j \bigcap (\bar{y}_j + \frac{\varepsilon}{2}\mathbb{B})\right] \times \left[\mathrm{cl}W \bigcap (\bar{w} + \frac{\varepsilon}{2}\mathbb{B})\right] \tag{8.25}$$

和一个均衡价格 $p^* \in E^* \setminus \{0\}$, 满足 (8.15) 和 (8.16), 以及

$$\langle p^*, u_i - x_i \rangle \geqslant -\varepsilon \parallel u_i - x_i \parallel, \quad \forall u_i \in \mathrm{cl}P_i(\bar{x}),\ i = 1, \cdots, n, \tag{8.26}$$

$$\langle p^*, v_j - y_j \rangle \leqslant \varepsilon \parallel v_j - y_j \parallel, \quad \forall v_j \in \mathrm{cl}S_j,\ j = 1, \cdots, m. \tag{8.27}$$

证明 由定理 8.5 中的关系 (8.13) 与 (8.14) 和命题 1.3 中的凸集 ε- 法向量的表示可直接得出. △

下一个定理建立了相当令人惊讶的结果: 基于类 Fréchet 法向量的光滑变分描述, 由近似的第二福利定理得到了一般的非凸经济学中边际价格的均衡含义. 事实上, 这样得到一个如前面的推论中凸型的扰动分散均衡, 但没有凸性假设, 并由一些非线性的价格取代 (8.26) 和 (8.27) 中的线性价格 p^*, 在特定意义上, 这些非线性价格是可微的, 其导数 (即变化率) 可在次优配置中任意接近 p^*.

定理 8.7(通过非线性价格的非凸经济学中的分散均衡)　对任意的 $\varepsilon > 0$, 下列断言成立:

(i) 设定理 8.5 中的所有假设均成立. 那么存在一个符合下面条件的次优三元组 (x, y, w) 满足 (8.25), 其中 w 为 (8.1) 中所定义, 存在满足 (8.15) 和 (8.16) 关系的一个边际价格 $p^* \in E^* \setminus \{0\}$, 以及存在商品空间 E 上实值函数, g_i, $i = 1, \cdots, n$ 和 h_j, $j = 1, \cdots, m+1$, 它们分别在 x_i, y_j 和 w 是 Fréchet 可微的, 并满足

$$\begin{cases} \|\nabla g_i(x_i) - p^*\| \leqslant \varepsilon, & i = 1, \cdots, n, \\ \|\nabla h_j(y_j) - p^*\| \leqslant \varepsilon, & j = 1, \cdots, m, \\ \|\nabla h_{m+1}(w) - p^*\| \leqslant \varepsilon, \end{cases} \tag{8.28}$$

还满足每个 $g_i, i = 1, \cdots, n$, 于 x_i 点在 $\mathrm{cl}P_i(\bar{x})$ 上实现全局最小, 每个 $h_j, j = 1, \cdots, m$, 于 y_j 点在 $\mathrm{cl}S_j$ 上实现全局最小, 且 h_{m+1} 于 w 点在 $\mathrm{cl}W$ 上实现全局最大.

(ii) 除了定理 8.5 中的假设, 假设 E 上存在一个在定理 1.30 中所考虑的 $\mathcal{S}$ 类的一个 $\mathcal{S}$- 光滑阻尼函数. 那么存在一个满足 (8.25) 的次优三元组 (x, y, w), 一个满足 (8.15) 和 (8.16) 边际价格 $p^* \in E^* \setminus \{0\}$, 以及满足 (8.28)$E$ 上的 $\mathcal{S}$ 光滑函数 g_i 和 h_j, 每个 g_i 唯一地于 x_i 点在 $\mathrm{cl}P_i(\bar{x})$ 上实现全局最小, 并且 $h_j, j = 1, \cdots, m$, 唯一地于 y_j 点在 $\mathrm{cl}S_j$ 上实现全局最大, 并且 h_{m+1} 唯一地于 w 点在 $\mathrm{cl}S_j$ 上实现全局最大. 此外, 如果 E 允许一个 Fréchet 光滑重赋范, 则可以选择 g_i 和 h_j 分别为凸和凹的.

证明　设 p^* 满足定理 8.5 的结论, 然后, 通过 (8.13)~(8.15), 可以找到 p_i^* 和 p_j^* 满足

$$\begin{cases} -p_i^* \in \widehat{N}(x_i; \mathrm{cl}P_i(\bar{x})), & \|p_i^* - p^*\| \leqslant \varepsilon, \quad i = 1, \cdots, n, \\ p_j^* \in \widehat{N}(y_i; \mathrm{cl}S_j), & \|p_j^* - p^*\| \leqslant \varepsilon, \quad j = 1, \cdots, m, \\ p_{j+1}^* \in \widehat{N}(w; \mathrm{cl}W), & \|p_{j+1}^* - p^*\| \leqslant \varepsilon \ . \end{cases} \tag{8.29}$$

现在应用定理 1.30 中的 Fréchet 法向量的光滑变分描述, 即得定理的所有结论. △

可以看到, 函数 g_i 和 h_j 发挥了定理 8.7 之前所讨论的非线性价格的作用, 它们确保了小扰动下非凸经济学中的分散凸类均衡.

8.2.2 第二福利定理的确切版本

下面导出经济 $\mathcal{E}$ 的 Pareto 和弱 Pareto 最优配置的点基必要最优条件, 这是由偏好、生产和净需求约束集的在最优配置点的基本法锥来表达的. 所取得的成果以

扩展的第二福利定理的确切版本给出. 其中在所考虑的优化配置下, 所有上面列出的经济集被赋予相同的边际价格, 从而提供了一个边际价格均衡.

确切的第二福利定理的证明是在前一小节建立的近似第二福利定理的基础上取极限. 如本书其他地方一样, 这个极限过程需要一些序列法紧性条件. 然而, 这里讨论的经济模型不同于所有以前的情形, 特别是不同于定理 2.22 里确切极点原理中的该条件. 所考虑的模型 $\mathcal{E}$ 的特别之处在于, 它只要求偏好、生产和净需求约束之中的一个具有这个性质, 而不是像其他地方那样要求除一个集合外全都具有这个性质. 经济框架 $\mathcal{E}$ 中确切极点原理的这样一个重要的改进主要是由于极值系统中集合 (8.18) 的分离结构.

定理 8.8(具有 Asplund 商品空间的扩展第二福利定理的确切形式) 设 $(\bar{x}, \bar{y})$ 为经济 $\mathcal{E}$ 中局部 Pareto(对应地, 弱 Pareto) 最优配置, 满足定理 8.5 中相应的假设, 其中 $\bar{w}$ 由 (8.2) 定义. 此外假设下述集合

$$\mathrm{cl}P_i(\bar{x}), \quad i = 1, \cdots, n; \quad \mathrm{cl}S_j, \quad j = 1, \cdots, m; \quad \mathrm{cl}W$$

中的一个在 $\bar{x}_i$, 或 $\bar{y}_j$, 或 $\bar{w}$ 点处是序列法紧的. 那么就存在一个非零价格 $p^* \in E^*$ 满足

$$-p^* \in N\big(\bar{x}_i; \mathrm{cl}P_i(\bar{x})\big), \quad i = 1, \cdots, n, \tag{8.30}$$

$$p^* \in N(\bar{y}_j; \mathrm{cl}S_j), \quad j = 1, \cdots, m, \tag{8.31}$$

$$p^* \in N(\bar{w}; \mathrm{cl}W). \tag{8.32}$$

证明 这里由定理 8.5 中的关系取极限得到证明. 选择任意序列 $\varepsilon_k \downarrow 0$ (当 $k \to \infty$ 时), 根据定理 8.5 可得序列 $(x_k, y_k, w_k, p_k^*) \in E \times E \times E \times E^*$ 满足

$$x_{ik} \in \mathrm{cl}P_i(\bar{x}) \cap (\bar{x}_i + \varepsilon_k \mathbb{B}), \quad i = 1, \cdots, n,$$

$$y_{jk} \in \mathrm{cl}S_j \cap (\bar{y}_j + \varepsilon_k \mathbb{B}), \quad j = 1, \cdots, m,$$

$$w_k = \sum_{i=1}^{n} x_{ik} - \sum_{j=1}^{m} y_{jk} \in \mathrm{cl}W \cap (\bar{w} + \varepsilon_k \mathbb{B}),$$

并且, 对于每个 $k \in \mathbb{N}$, 都有 $\varepsilon = \varepsilon_k$ 时的对偶关系 (8.13)~(8.16). 显然 $k \to \infty$ 时 $(x_k, y_k, w_k) \to (\bar{x}, \bar{y}, \bar{w})$. 既然 E 是 Asplund, 且由 (8.16) 知价格 p_k^* 一致有界, 所以存在 $p^* \in E^*$ 使得序列 $\{p_k^*\}$ 在 E^* 的弱 $*$ 拓扑下收敛于 p^*. 当 $k \to \infty$ 时在 (8.13)~(8.15) 中取极限, 并由基本法锥的结构, 就得到 (8.30)~(8.32) 中的所有关系.

剩下要证明的是, 如果集合 $P_i(\bar{x})$, $\mathrm{cl}S_j$, $\mathrm{cl}W$ 中的一个在对应点是 SNC 的, 那么 $p^* \neq 0$. 若不然, 设 $p^* = 0$. 明确起见假定集合 $\mathrm{cl}W$ 在 $\bar{w}$ 是 SNC 的. 则由 (8.15) 可得序列 $e_k^* \in E$ 满足

$$p_k^* - \varepsilon_k e_k^* \in \widehat{N}(w_k; \mathrm{cl}W), \quad \|e_k^*\| = 1, \quad \forall k \in \mathbb{N}. \tag{8.33}$$

很明显, 当 $k \to \infty$ 时, $p_k^* - \varepsilon_k e_k^* \xrightarrow{w^*} 0$. 然后通过 SNC 集合的定义 1.20, 从 (8.33) 得 $\|p_k^* - \varepsilon_k e_k^*\| \to 0$. 因此当 $k \to \infty$ 时, $\|p_k^*\| \to 0$. 很明显后者与 (8.16) 左边关于 p_k^* 的不等式矛盾. 这样就得到 $p^* \neq 0$, 定理得证. △

现在讨论定理 8.8 的一些有用的推论. 首先考虑 $\mathcal{E}$ 的一个特殊情形, 其中净需求约束集 W 对于某 $\omega \in \mathrm{cl}W$ 具有锥表示

$$W = \omega + \Gamma. \tag{8.34}$$

如果 E 是一个包含闭正锥 E_+ 的有序商品空间 (参看下一节), 那么在 $\Gamma := -E_+$ 时 (8.34) 对应于所谓的商品隐含自由处置. 现在考虑一个更一般的例子: Γ 是 E 中任意的凸锥, 不具有有序结构, 在这个情形下可证 (8.32) 蕴涵如下互补松弛性条件, 这在经济上可以理解为边际价格的零值过剩需求.

推论 8.9(过剩需求条件)　除了定理 8.8 的假设, 假设 W 由 (8.34) 给出, Γ 为商品空间 E 的非空凸子锥. 那么存在一个非零价格 $p^* \in E^*$ 满足 (8.30) 和 (8.31), 并且

$$\left\langle p^*, \sum_{i=1}^n \bar{x}_i - \sum_{j=1}^m \bar{y}_j - \omega \right\rangle = 0. \tag{8.35}$$

证明　为了证明 (8.35), 注意到

$$\langle p^*, \bar{w} - \omega \rangle \geqslant \langle p^*, w - \omega \rangle, \quad \forall w \in \mathrm{cl}W, \tag{8.36}$$

这是由于 (8.32)、(8.34) 和凸集的法锥表示. 因此 $\langle p^*, \bar{w} - \omega \rangle \geqslant 0$. 另一方面. 由于 Γ 的锥结构, 取

$$2(\bar{w} - \omega) \in W - \omega = \Gamma.$$

通过 (8.36) 得到 $\langle p^*, \bar{w} - \omega \rangle \leqslant 0$, 这就证明了 (8.35) 并且完成了该推论的证明. △

在具有凸的偏好和生产集的经济情况下, 定理 8.8 的 (8.30) 和 (8.31) 关系归结为福利经济学中第二基本定理的经典的消费者支出最小化和公司利润最大化条件. 然而, 这里可以在本质上改进加在经济 $\mathcal{E}$ 中凸集的非空内部条件. 事实上, 从定理 1.21 可知, 第二福利定理的扩展所要求的 SNC 性质, 对具有非空相对内部的凸集, 等价于该集合的有限余维数性质. 此外, Asplund 空间的凸集即使相对内部是空的仍可能是 SNC 的; 见例 3.6 和备注 1.27 的讨论. 因此, 定理 8.8 以下的推论, 在 Pareto 和弱 Pareto 最优两种情况下, 对具有 Asplund 商品空间的凸经济的经典第二福利定理都是一个长足的改进.

推论 8.10(凸经济的改进的第二福利定理)　除了定理 8.8 的假设, 假设所有的偏好和生产集 $P_i(\bar{x})$, $i = 1, \cdots, n$ 和 S_j, $j = 1, \cdots, m$ 是凸的. 那么存在一个非

零价格 $p^* \in E^*$ 满足 (8.32), 并且

$$\text{对任意 } i=1,\cdots,n, \quad x_i \text{ 在 } \mathrm{cl}P_i(\bar{x}) \text{ 上极小化} \langle p^*, x_i \rangle;$$

$$\text{对任意 } j=1,\cdots,m, \quad y_j \text{ 在} \mathrm{cl}S_j \text{ 上极大化} \langle p^*, y_j \rangle.$$

证明 由凸集的法锥表示, 这个结论可从 (8.30) 和 (8.31) 直接得到. △

注 8.11(非凸均衡) 如上述推论所示, 定理 8.8 的假设证明了在福利经济学凸模型的 Pareto 和弱 Pareto 的最优配置中的分散价格均衡. 与定理 8.7 的近似/次优情形不同, 一般不能通过非线性价格提供点基关系 (8.30) 和 (8.31) 的分散均衡解释. 然而, 本节的结果使得可以在非凸模型中把定理 8.8 中的一阶必要最优条件里给出的边际价格均衡看成凸类型的分散均衡的极限情形, 这可以通过使用非线性价格得以实现.

8.3 有序商品空间的非凸经济

本节研究当商品空间 E 为一个有序的 Banach 空间时的这种特殊的福利经济模型 $\mathcal{E}$. 研究的目标为:

(i) 在定理 8.8 中给出的 (确切) 扩展第二福利定理框架中找到确保 (一般为非凸) 经济 $\mathcal{E}$ 的 Pareto 和弱 Pareto 最优配置的边际价格为正的有效条件.

(ii) 在有序商品空间情形下并且不加净需求规范条件, 得出关于强 Pareto 最优配置的近似和确切的第二福利定理的新版本.

这些目标在下述两节中达成. 注意这里没有对所论的商品空间加任何格结构.

8.3.1 正的边际价格

设 E 为有序的 Banach 空间, 其中闭正锥为

$$E_+ := \{e \in E \,|\, e \geqslant 0\}.$$

依照经济学文献传统标记法, 其中 (标准) 偏序关系记作 $\geqslant$. 相应的对偶正锥 E_+^* 是序空间 E^* 的闭正锥, 它具有表示:

$$E_+^* := \{e^* \in E^* \,|\, e^* \geqslant 0\} = \{e^* \in E^* \,|\, \langle e^*, e \rangle \geqslant 0, \forall e \in E_+\},$$

其中 E^* 上的序是通过 E 上的 $\geqslant$ 而导出的.

边际价格的正性条件基于下面的引理, 它确保了有序 Banach 空间中的闭子集的基本法锥的正性, 其本身也具有独立的意义.

引理 8.12(有序空间的基本法锥的正性) 设 E 为有序 Banach 空间, Ω 为 E 的非空闭子集并满足

$$\Omega - E_+ \subset \Omega, \tag{8.37}$$

则有包含关系

$$N(\bar{e};\Omega) \subset E_+^*, \quad \forall \bar{e} \in \Omega. \tag{8.38}$$

证明 取 $e^* \in N(\bar{e};\Omega)$, 其中 Ω 满足 (8.37). 根据基本法锥的定义, 存在序列

$$\varepsilon_k \downarrow 0, \quad e_k \xrightarrow{\Omega} \bar{e}, \quad e_k^* \xrightarrow{w^*} e^* \quad (k \to \infty),$$

其中对所有的 $k \in \mathbb{N}$ 有 $e_k^* \in \widehat{N}_{\varepsilon_k}(e_k;\Omega)$. 由 (8.37) 和 ε- 法锥的单调性, 当 $e \in \Omega_2 \subset \Omega_1$ 和 $\varepsilon \geqslant 0$ 时, $\widehat{N}_\varepsilon(e;\Omega_1) \subset \widehat{N}_\varepsilon(e;\Omega_2)$. 因此对所有的 $k \in \mathbb{N}$, 都有 $e_k^* \in \widehat{N}_{\varepsilon_k}(e_k;\Omega - E_+)$. 固定 $k \in \mathbb{N}$ 并任取 $\gamma > 0$. 由 ε- 法锥的定义可找到 $\eta_k > 0$ 满足

$$\langle e_k^*, e - e_k\rangle \leqslant (\varepsilon_k + \gamma)\|e - e_k\|, \quad \forall e \in (e_k + \eta_k\mathbb{B}) \cap (\Omega - E_+). \tag{8.39}$$

很容易看出

$$e_k - \eta_k u \in (e_k + \eta_k\mathbb{B}) \cap (\Omega - E_+), \quad \forall u \in E_+ \cap \mathbb{B}.$$

把 $e := e_k - \eta_k u$ 代入 (8.39) 得

$$\langle e_k^*, -u\rangle \leqslant (\varepsilon_k + \gamma)\|u\| \leqslant \varepsilon_k + \gamma, \quad \forall u \in E_+ \cap \mathbb{B}, \quad k \in \mathbb{N}.$$

对该不等式取极限, 并注意到 $e_k^* \xrightarrow{w^*} e^*$ $(k \to \infty)$, 则得

$$\langle e^*, -u\rangle \leqslant \gamma, \quad \forall u \in E_+ \cap \mathbb{B}.$$

由于 $\gamma > 0$ 是任意的, 这意味着 $e^* \in E_+^*$. 这就给出 (8.38) 并完成引理证明. △

注意 (8.37) 与经济模型中的自由处置条件相关. 下一定理包括加在偏好, 或者生产, 或者净需求约束集合的该类假设, 它确保在有序 Asplund 商品空间的框架中, 扩展第二福利定理中的价格正性 $p^* \in E_+^* \setminus \{0\}$.

定理 8.13(Pareto 和弱 Pareto 最优配置的正价格) 设 $(\bar{x}, \bar{y})$ 为经济 $\mathcal{E}$ 的局部 Pareto(相应地, 弱 Pareto) 最优配置. 除了定理 8.8 的相应假设, 假设 E 是一个有序空间且下列条件之一成立:

(a) 存在 $i \in \{1, \cdots, n\}$ 使得第 i 个消费者在 $\bar{x}$ 点满足期望条件, 即

$$\mathrm{cl}P_i(\bar{x}) + E_+ \subset \mathrm{cl}P_i(\bar{x}).$$

(b) 存在 $j \in \{1, \cdots, m\}$ 使得第 j 个企业满足自由处置条件, 即

$$\mathrm{cl}S_j - E_+ \subset \mathrm{cl}S_j.$$

(c) 净需求约束集合 W 具有隐性商品自由处置性质, 即

$$\mathrm{cl}W - E_+ \subset \mathrm{cl}W.$$

那么存在一个正的边际价格 $p^* \in E_+^* \setminus \{0\}$ 满足通过基本法锥给出的关系 (8.30)~(8.32).

证明 由定理 8.8 中的关系 (8.31) 和 (8.32), 在 (b) 和 (c) 情形中的边际价格的正性 $p^* \in E_+^*$ 可以直接从引理 8.12 得来. 情形 (a) 可由 (8.30) 和下述性质归结为引理 8.12; 该性质为

$$N(\bar{e}; \Omega) = -N(-\bar{e}; -\Omega), \quad \forall \Omega \subset E, \bar{e} \subset \Omega.$$

此性质在任何 Banach 空间皆成立, 这可通过定义来直接验证. △

可以看到, 若 $\mathrm{int}E_+ \neq \varnothing$, 则 (a)~(c) 的每一个条件都蕴涵着相应集合 $P_i(\bar{x})$, $\mathrm{cl}S_j$ 和 $\mathrm{cl}W$ 的上图 Lipschitz 性质, 由上述讨论, 对 E 的正锥非空内部要求也保证了定理 8.8 的规范和法紧性条件, 因此在定理 8.13 中存在一个正边际价格 $p^* \in E_+^* \setminus \{0\}$.

8.3.2 强 Pareto 最优的改进结果

可以看到, 净需求规范条件 (NDQ 和 NDWQ) 在上述关于 Pareto 和弱 Pareto 最优配置的扩展第二福利定理的证明中发挥了主要的作用. 事实上, 它们允许把相应的 Pareto 最优归结为集合系统的极值点, 从而可以运用极点原理. 在有序商品空间 E 的情形, 这些规范条件相关于 E 的正锥的非空内部条件 $\mathrm{int}E_+ \neq \varnothing$. 当然, 因为强 Pareto 最优是一个有更多限制的概念, 关于强 Pareto 最优配置以上所有扩展的第二福利定理都成立.

这一小节证明, 在有序商品空间中, 凸与非凸经济的强 Pareto 最优配置根本不需要净需求规范条件, 其中在许多配置中 $E_+ \neq \varnothing$, 这对理论和应用都很重要. 强 Pareto 最优性这个要求恰巧可以把相应的最优配置归结为没有任何规范条件的集合系统的局部极值点. 因此又可以使用极点原理这个变分分析的主要工具.

现在回顾一下生成锥的概念. 称闭正锥 E_+ 对 E 是生成的, 如果该空间可表示为 $E = E_+ - E_+$. 由正生成锥所定序的 Banach 空间类很大, 特别包括所有 Banach 格 (或赋范完备 Riesz 空间), 它们的生成正锥的内部一般都是空的.

下面的结果为具有有序 Asplund 商品空间的 (一般非凸) 经济的强 Pareto 最优配置建立了第二福利定理的几个版本. 它包含第二福利定理近似和确切两个形式, 包括期望/自由处置类型假设下的边际价格的正性. 值得注意的是, 定理的前两个陈述要求正锥是生成的, 而第三个陈述提供了经济上的替代假设, 它保证了在更一般的有序空间得出同样的结论.

定理 8.14(强 Pareto 最优配置第二福利定理) 设 $(\bar{x}, \bar{y})$ 为一个具有有序 Asplund 商品空间的经济 $\mathcal{E}$ 的局部强 Pareto 最优配置, 设集合 S_j, W 分别在 $\bar{y}_j$, $\bar{w}$ 附近是局部闭的. 那么下列成立:

(i) 假设闭正锥 E_+ 是生成的, 或者经济显示出商品的隐含自由处置

$$W - E_+ \subset W \tag{8.40}$$

或者对于某 $j \in \{1, \cdots, m\}$ 自由处置生产条件

$$S_j - E_+ \subset S_j \tag{8.41}$$

成立, 或者 $n > 1$, 且存在一个消费者 $i_0 \in \{1, \cdots, n\}$ 使得 $P_{i_0}(\bar{x}) \neq \varnothing$, 且对于某 $i \in \{1, \cdots, n\} \backslash \{i_0\}$, 有期望条件

$$\text{cl} P_i(\bar{x}) + E_+ \subset \text{cl} P_i(\bar{x}). \tag{8.42}$$

那么对于每一个 $\varepsilon > 0$, 都存在一个次优三元组:

$$(x, y, w) \in \prod_{i=1}^{n} \left[\text{cl} P_i(\bar{x}) \cap \left(\bar{x}_i + \frac{\varepsilon}{2}\mathbb{B}\right)\right] \times \prod_{j=1}^{m} \left[S_j \cap \left(\bar{y}_j + \frac{\varepsilon}{2}\mathbb{B}\right)\right] \times \left[W \cap \left(\bar{w} + \frac{\varepsilon}{2}\mathbb{B}\right)\right],$$

其中综合商品 w 定义于 (8.1), 公共边际价格 $p^* \in E^*$ 满足关系 (8.13)~(8.16).

(ii) 在 (i) 中假设的基础上, 如果下列集合中的一个

$$\text{cl} P_i(\bar{x}), \quad i = 1, \cdots, n, \quad S_j, \quad j = 1, \cdots, m, \quad W$$

在相应点是 SNC 的, 那么存在一个满足点基关系 (8.30)~(8.32) 的正边际价格 $p^* \in E^* \backslash \{0\}$.

(iii) 如果不假设 E_+ 是一个生成锥, 而假设 $E_+ \neq \{0\}$, 且集合 W, 对于 $j = 1, \cdots, m$ 的集合 S_j, 对于 $i = 1, \cdots, n$ 的集合 $P_i(\bar{x})$ 中至少有两个满足 (8.40)~(8.42) 的相应条件, 则 (i) 和 (ii) 中所有结论也都成立.

证明 考虑 (8.17) 和 (8.18) 中定义的集合系统 $\{\Omega_1, \Omega_2\}$, 其中 (8.17) 中 S_j 和 W 的闭包运算被省略, 因为这些集合在相关点附近是局部闭的. 取 $\mathcal{E}$ 的一个强 Pareto 局部最优 $(\bar{x}, \bar{y})$, 下面证明, 如果 (i) 或 (iii) 中的假设成立, 那么 $(\bar{x}, \bar{y}, \bar{w}) \in \Omega_1 \cap \Omega_2$ 是 $\{\Omega_1, \Omega_2\}$ 的局部极值点. 这样, 这些假设就取代了定理 8.5 中在证明 Pareto 和弱 Pareto 最优配置时相应的净需求规范条件.

首先在所给的假设下考虑 (i). 确定起见假定 (8.40) 成立; 其他两种情况类似.

容易看到, $\bar{w}$ 是 W 的一个边界点; 否则, 由于对于某 $i \in \{1, \cdots, n\}$, $P_i(\bar{x}) \neq \varnothing$, 这与 $(\bar{x}, \bar{y})$ 的 Pareto 最优相矛盾. 因此, 可找到 E 中的序列 $e_k \to 0$, 对于所有的

$k \in \mathbb{N}$ 都满足于 $\bar{w} + e_k \notin W$. 由经典的 Krein-Šmulian 定理 (例如, 参看 Abramovich 和 Aliprantis[1] 中的证明、讨论和参考文献), 在任何由闭生成锥定序的 Banach 空间 E 中, 存在一个常数 $M > 0$, 使得对每个 $e \in E$ 都有正矢量

$$u, v \in E_+ \text{ 满足 } e = u - v, \text{且} \max\{\|u\|, \|v\|\} \leqslant M\|e\|.$$

由此可找到满足 $e_k = u_k - v_k$ 的序列 $u_k \xrightarrow{E_+} 0$ 和 $v_k \xrightarrow{E_+} 0$. 由于 $v_k \in E_+$, $W - E_+ \subset W$, 所以有

$$\bar{w} + u_k \notin W, \tag{8.43}$$

其中, 当 $k \to \infty$ 时, $u_k \xrightarrow{E_+} 0$.

现由局部 Pareto 最优配置 $(\bar{x}, \bar{y})$ 的定义取邻域 $O \subset E^{n+m}$, 下证定理 8.5 证明中的极值条件 (8.19) 对于所有的 $k \in \mathbb{N}$ 沿序列 $a_k := (0, \cdots, 0, u_k) \in E^{n+m+1}$ 和邻域 $U := O \times E$ 成立. 这就证明了所论系统 $\{\Omega_1, \Omega_2\}$ 在 $(\bar{x}, \bar{y}, \bar{w})$ 点的局部极值性.

假设对某 $k \in \mathbb{N}$ 条件 (8.19) 不成立. 取 $(x_k, y_k, w_k) \in \Omega_1$ 满足 $(x_k, y_k) \in O$, $(x_k, y_k, w_k - u_k) \in \Omega_2$. 通过 $u_k \in E_+$ 和隐式自由处置假设 (8.40), 则 (x_k, y_k) 的分量满足

$$\sum_{i=1}^{n} x_{ik} - \sum_{j=1}^{m} y_{jk} = w_k - u_k \in W - E_+ \subset W, \tag{8.44}$$

这意味着 (x_k, y_k) 是属于取定的 $(\bar{x}, \bar{y})$ 邻域的经济 $\mathcal{E}$ 的可行配置. 由于 $(\bar{x}, \bar{y})$ 是 $\mathcal{E}$ 的强 Pareto 最优配置, 对于所有的大 $k \in \mathbb{N}$, 有 $(x_k, y_k) = (\bar{x}, \bar{y})$. 由此就得到

$$\begin{aligned} \bar{w} + u_k &= \sum_{i=1}^{n} \bar{x}_i - \sum_{j=1}^{m} \bar{y}_j + u_k = \sum_{i=1}^{n} x_{ik} - \sum_{j=1}^{m} y_{jk} + u_k \\ &= (w_k - u_k) + u_k = w_k \in W, \end{aligned}$$

这与 (8.43) 矛盾, 这样就证明了情形 (i) 时 $\{\Omega_1, \Omega_2\}$ 在 $(\bar{x}, \bar{y}, \bar{w})$ 点的局部极值性.

接下来证明, 在定理 (iii) 的情形, 当正锥 E_+ 可能非生成时, 极值条件 (8.19) 也成立. 明确起见, 假设隐含自由处置条件 (8.40) 成立且一个生产集 (比如 S_1) 满足自由处置条件 (8.41). 选择一个对于所有 $k \in \mathbb{N}$ 满足 $u_k \neq 0$ 的序列 $u_k \xrightarrow{E_+} 0$; 由于 $E_+ \neq \{0\}$, 这总是可能的. 再选择 $a_k := (0, \cdots, 0, u_k) \in X$, 下面按照该序列检验 (8.19). 否则, 重复上面的论证, 可以找到满足 (8.44) 的 $(x_k, y_k, w_k) \in \Omega_2 \cap U$. 则对所有的大 $k \in \mathbb{N}$ 有 $(x_k, y_k) = (\bar{x}, \bar{y})$, 这是因为 $(\bar{x}, \bar{y})$ 是 $\mathcal{E}$ 的局部强 Pareto 最优配置. 在这个情形下, 从 (8.44) 可以得出, 对于足够大的 $k \in \mathbb{N}$,

$$\sum_{i=1}^{n} x_{ik} - (y_{1k} - u_k) - \sum_{j=2}^{m} y_{jk} = w_k \in W \tag{8.45}$$

通过 $j=1$ 时的自由处置条件 (8.41), 得到 $y_{1k}-u_k\in S_1$, 因此 (8.45) 确保 $(x_k, y_k-u_k, 0, \cdots)$ 属于强 Pareto 局部最优 $(\bar{x},\bar{y})$ 取定邻域的 $\mathcal{E}$ 的可行配置. 因此

$$y_{1k}-u_k=\bar{y}_1-u_k=\bar{y}_1,$$

即对于所有大的 $k\in\mathbb{N}$, $u_k=0$. 这一矛盾证实了 (iii) 中的假设下系统 $\{\Omega_1,\Omega_2\}$ 在 $(\bar{x},\bar{y},\bar{w})$ 点的局部极值性.

把定理 2.20 的极点原理运用到该集合系统中, 即得定理 8.5 中罗列的近似第二福利定理的结论, 但是在 (i) 或 (iii) 的假设下和强 Pareto 最优配置情形中, 这里没有假设 SNC 性质. 如果最后另外添加 (ii) 中的 SNC 假设, 则如定理 8.8 的证明, 可取极限而获得扩展第二福利定理的确切关系 (8.30)~(8.32). (8.40)~(8.42) 中价格的正性从引理 8.12 中得来, 如定理 8.13 的证明. 证毕. △

8.2 节建立的定理 8.5 和定理 8.8 的推论, 包括均衡解释, 不取决于净需求规范条件, 在定理 8.14 的框架中, 它们对于强 Pareto 最优配置也成立.

注 8.15(强 Pareto 最优配置的修改概念和结果) Glenn Malcolm(个人交流) 最近观察到, 关于强 Pareto 最优配置在 8.3.2 节所获得的扩展第二福利定理的结果, 不需修改即对强 Pareto 最优性的一个修改版本同样适用, 而这对经济应用可能更具吸引力. (局部) 强 Pareto 最优配置的这个修改版本与定义 8.2(iii) 的唯一的区别如下: 8.2(iii) 要求对任意可行配置 $(x,y)\neq(\bar{x},\bar{y})$, 对于某 $i\in\{1,\cdots,n\}$, 条件 $x_i\notin \mathrm{cl}P_i(\bar{x})$ 成立, 现在要求对那些满足 $x\neq\bar{x}$ 的 (x,y)(局部) 满足这一条件.

这一修改允许把具有不同生产计划和相关资产的可行配置考虑在内, 这似乎具有重大的经济意义.

读者可以验证, 只需对上述定理 8.14 的证明作少许改动, 即可对该修改的强 Pareto 最优建立近似和确切的扩展第二福利定理, 其中的假设与定理 8.14 中的 (i)~(iii) 完全一样, 而不需要净需求规范条件.

事实上, 考虑集合

$$\Sigma := W+\sum_{j=1}^{m} S_j\cap(\bar{y}_j+\nu\mathbb{B}),$$

这里 $\nu>0$ 足够小. 容易验证, 商品

$$\bar{w}+\sum_{j=1}^{m}\bar{y}_j$$

给出了集合 Σ 的一个边界点. 通过使用 Krein-Šmulian 定理和用 $\bar{w}+\sum\limits_{j=1}^{m}\bar{y}_j\notin\Sigma$, 这里当 $k\to\infty$ 时 $u_k\xrightarrow{E_+}0$, 替代条件 (8.43), 可以同样地进行定理 8.14(i) 的证明.

基于强 Pareto 最优配置修改的定义, 如定理 8.14, 这就能证明三元组 $(\bar{x}, \bar{y}, \bar{w})$ 是集合系统 $\{\Omega_1, \Omega_2\}$ 的一个局部极值点. 然后可以用类似推理来证明修改情形下对应于定理 8.14 的 (ii) 和 (iii) 的部分.

8.4 抽象版本和进一步扩展

本章最后一节包含一些与以上研究相比更一般框架下的经济学建模的结果和讨论. 首先, 提出 8.1 节描述的经济 $\mathcal{E}$ 的扩展第二福利定理的抽象预法锥的版本, 这里没有要求商品空间的 Asplund 结构. 最后一节关注具有公共商品的福利经济学模型, 以及一些包括具有公众环境与直接分配模型的进一步扩展.

8.4.1 第二福利定理的抽象版本

在 8.1.1 节中已指出, 上述考虑的福利经济模型 $\mathcal{E}$ 在任意线性拓扑空间中都是有意义的. 但 8.2~8.3 节所得出的扩展第二福利定理结果, 是在具有 Asplund 商品空间的经济中以 Fréchet 和基本法向量来陈述和证明的. 分析 8.2~8.3 节给出的结果的证明可以看出, 使用类 Fréchet 结构, 或者商品空间的 Asplund 结构, 或者两者兼而有之的需求主要来自以下两方面:

(i) 使用由一般闭集的类 Fréchet 法向量描述的定理 2.20 的极点原理. 由于该原理是 Asplund 空间的刻画, 这里的 Asplund 性质是不可避免的.

(ii) 涉及正价格的定理 8.13 和定理 8.14 是基于推论 8.12 的. 虽然该推论在任意 Banach 空间上成立, 但是似乎大量地使用了一些 Fréchet 和基本法锥的特殊性质. 此外, 定理 8.7 关于非凸经济以非线性价格给出的扩展第二福利定理的分散均衡描述以及注 8.11 的相关讨论, 无疑取决于确保它们变分表示的广义法向量的类 Fréchet 结构. 如上所述, 定理 8.7(ii) 列出的所论空间的特殊几何性质其实蕴涵着商品空间的 Asplund 性质.

对非 Fréchet 类型的广义法向量而言, 似乎不大可能获得类似的变分描述, 而这对边际价格均衡的上述分散解释至关重要. 另一方面, 基于极点原理的 (i) 中的主要结果, 在非 Asplund 空间有对应的版本, 这是通过相应的预法锥和法锥结构给出的; 见 2.5 节, 那里导出了极点原理的某些抽象/公理版本. 本节的目标是澄清需要什么样的预法锥和法锥结构的附加假设, 来保证 8.2 节和 8.3 节建立的近似/确切第二福利定理的类似的抽象版本成立.

现在从 Pareto、弱 Pareto、强 Pareto 最优配置的第二福利定理的近似版本入手. 注意定义 8.3 的净需求规范条件与定理 8.14 罗列的自由处置/期望条件不涉及广义法向量. 除了预法锥结构定义 2.41 中的 (H) 和命题 2.42、命题 2.43 中蕴涵着 (H) 的预次微分结构的性质, 还需要广义法向量的其他性质, 它们足以在相应的

Banach 空间背景中获得定义 8.5 和定义 8.14(ii) 和 (iii) 对应的抽象版本.

接下来, 除了 (H), 假设 Banach 空间 X 上法锥的如下性质, 它们对任何合理的预法锥结构 $\widehat{\mathcal{N}}(\cdot;\Omega)$ 无疑都是成立的.

(H1) 如果 $\Omega \subset X$ 为 X 的线性子空间, 且 $\bar{x} \in \Omega$, 那么

$$\widehat{\mathcal{N}}(\bar{x};\Omega) = \Omega^{\perp} := \{x^* \in X^* \big| \langle x^*, x\rangle = 0, \quad \forall x \in \Omega\}$$

是一个与 Ω 正交的子空间.

(H2) 对于所有 X 的闭子集 Ω_1 和 Ω_2 和 $\bar{x}_i \in \Omega_i$, $i = 1, 2$, 有

$$\widehat{\mathcal{N}}\big((\bar{x}_1,\bar{x}_2);\Omega_1 \times \Omega_2\big) \subset \widehat{\mathcal{N}}(\bar{x}_1;\Omega_1) \times \widehat{\mathcal{N}}(\bar{x}_2;\Omega_2).$$

注意, 通过命题 1.2, 乘积性质 (H2) 对于 Fréchet 法锥成立且为等式. 另外, 对由次微分生成的预法锥结构 $\widehat{\mathcal{N}}(\cdot;\Omega) = \widehat{\mathcal{D}}\delta(\cdot;\Omega)$, (H2) 总是可由 2.5.1 小节预次微分 $\widehat{\mathcal{D}}$ 的性质 (S3) 导出; 见命题 2.42 的证明.

下一个定理提供了在定理 8.5 中建立的关于 Pareto 和弱 Pareto 最优配置的近似第二福利定理的一个抽象版本.

定理 8.16(Pareto 和弱 Pareto 最优配置的近似第二福利定理的抽象版本)　设 $\mathcal{E}$ 为一个具有 Banach 商品空间 E 的经济, 设 $X := E^{n+m+1}$, 且 $\widehat{\mathcal{N}}$ 为具有性质 (H1) 和 (H2) 的 X 的一个预法锥结构. 考虑 $\mathcal{E}$ 的 Pareto(对应地, 弱 Pareto) 最优配置 $(\bar{x},\bar{y})$, 其中 $\bar{w}$ 由 (8.2) 定义. 假定净需求规范条件 (对应地, 净需求弱规范条件) 在 $(\bar{x},\bar{y})$ 成立. 那么对每个 $\varepsilon > 0$, 都存在一个次优三元组

$$(x,y,w) \in \prod_{i=1}^{n} \mathrm{cl}P_i(\bar{x}) \times \prod_{j=1}^{m} \mathrm{cl}S_j \times \mathrm{cl}W$$

(其中总商品 w 由 (8.1) 定义) 和共同边际价格 $p^* \in E^* \setminus \{0\}$ 满足 (8.13)~(8.16), 其中用 $\widehat{\mathcal{N}}$ 取代 $\widehat{N}$.

证明　沿用定理 8.5 的证明过程, 其中使用定理 2.51(i) 中近似极点原理的抽象版本 (它对任何预法锥结构都成立), 以及完成这一过程所需的性质 (H1) 和 (H2). △

由于凸分析法锥总满足定理 8.16 的假设 (H)、(H1) 和 (H2), 关于凸经济扰动分散均衡的推论 8.6 所有结论在任意 Banach 空间中成立. 如上所述, 这与关于非凸经济的定理 8.7 是不一样的.

对于具有有序商品空间经济的强 Pareto 最优配置, 第二福利定理的抽象近似版本不需要净需求规范条件, 见以下定理.

定理 8.17(强 Pareto 最优配置近似第二福利定理的抽象版本)　设 $(\bar{x},\bar{y})$ 为具有有序 Banach 空间 E 的经济 $\mathcal{E}$ 的一个局部强 Pareto 最优配置, 设集合 S_j 和 W

分别在 $\bar{y}_j$ 和 $\bar{w}$ 局部闭, 并且 $\widehat{\mathcal{N}}$ 为 $X = E^{n+m+1}$ 上满足性质 (H1) 和 (H2) 的预法锥结构. 假设如下条件之一成立:

(a)E_+ 是生成性的, 并且自由处置/期望性假设 (8.40)~(8.42) 中有一个满足;

(b)$\mathrm{int}\, E_+ \neq \varnothing$, 且集合 W, 对于 $j = 1, \cdots, m$ 的 S_j, 对于 $i = 1, \cdots, n$ 的 $P_i(\bar{x})$ 中至少两个满足相应的 (8.40)~(8.42) 中的自由处置/期望性假设.

那么对于每个 $\varepsilon > 0$, 都存在一个次优三元组

$$(x, y, w) \in \prod_{i=1}^{n}\left[\mathrm{cl}P_i(\bar{x})\bigcap\left(\bar{x}_i + \frac{\varepsilon}{2}\mathbb{B}\right)\right] \times \prod_{j=1}^{m}\left[S_j\bigcap\left(\bar{y}_j + \frac{\varepsilon}{2}\mathbb{B}\right)\right] \times \left[W\bigcap\left(\bar{w} + \frac{\varepsilon}{2}\mathbb{B}\right)\right]$$

(其中总商品 w 由 (8.1) 定义) 和共同边际价格 $p^* \in E^*$ 满足关系 (8.13)~(8.16), 其中 $\widehat{N}$ 被 $\widehat{\mathcal{N}}$ 取代.

证明 沿着定理 8.14(i, iii) 的证明并使用来自定理 2.51(i) 的抽象极点原理, 定理即可得证. △

接下来, 通过对相应的近似版本取极限, 得出具有 Banach 商品空间的非凸经济关于 Pareto、弱 Pareto 和强 Pareto 最优配置的第二福利定理的确切版本. 为此需要使用定义 2.50 引入的抽象序列法紧性条件 ($\widehat{\mathcal{N}}$-SNC). 这个条件依赖于给定的预法锥结构 $\widehat{\mathcal{N}}$. 当 $\widehat{\mathcal{N}} = \widehat{N}$ (即 Fréchet 法锥) 时, 对 Asplund 空间的闭子集它归结为基本的 SNC 性质. 注意, 虽然下面第二福利定理的抽象扩展需要序列法紧性, 但边际价格关系却是在一般意义上通过 Banach 空间上法锥结构的拓扑 (网) 极限来表达的. 这无疑是个很大的优点.

这里首先给出一个 Pareto 和弱 Pareto 最优配置的确切第二福利定理的抽象扩展, 它使用定义 8.3 中的净需求规范条件, 推广了定理 8.8 的相应结果. 如定理 8.8, 下面定理中的 $\widehat{\mathcal{N}}$-SNC 条件只需对偏好、生产和净需求约束集其中的一个成立即可, 这与定理 2.51(ii) 的一般确切极点原理不同, 在那里要求在极值系统中除了一个之外所有集合都满足该条件.

定理 8.18(Pareto 和弱 Pareto 最优配置的确切第二福利定理的抽象版本) 设 $(\bar{x}, \bar{y})$ 为经济 $\mathcal{E}$ 的局部 Pareto(对应地, 弱 Pareto) 最优配置, 并设定理 8.16 的相应假设都成立. 取 X 上的一个预法锥结构 $\widehat{\mathcal{N}}$, 假设集合

$$\mathrm{cl}\, P_i(\bar{x}), \quad i = 1, \cdots, n, \quad \mathrm{cl}\, S_j, \quad j = 1, \cdots, m, \quad \mathrm{cl}\, W$$

中的一个在相应点为 $\widehat{\mathcal{N}}$-SNC 的. 那么存在一个非零价格 $p^* \in E^*$ 满足关系

$$-p^* \in \overline{\mathcal{N}}\big(\bar{x}_i; \mathrm{cl}P_i(\bar{x})\big), \quad i = 1, \cdots, n, \tag{8.46}$$

$$p^* \in \overline{\mathcal{N}}(\bar{y}_j; \mathrm{cl}S_j), \quad j = 1, \cdots, m, \tag{8.47}$$

$$p^* \in \overline{N}(\bar{w}; \mathrm{cl}W), \tag{8.48}$$

其中 $\overline{N}$ 代表 $\widehat{N}$ 生成的拓扑法锥结构 (2.67). 此外, 拓扑结构 $\overline{N}$ 在 (8.46)~(8.48) 中可以被 (2.66) 中由 $\widehat{N}$ 生成的序列法锥结构 N 替换, 如果闭对偶球 $\mathbb{B}^* \subset E^*$ 为弱 * 列紧的.

证明　与定理 8.8 的证明相似, 取任意数列 $\varepsilon_k \downarrow 0\ (k \to \infty)$. 根据定理 8.16 的抽象近似版本, 找出序列 $(x_k, y_k, w_k, p_k^*) \in E \times E \times E \times E^*$ 满足

$$x_{ik} \in \mathrm{cl}P_i(\bar{x}) \cap (\bar{x}_i + \varepsilon_k \mathbb{B}), \quad i = 1, \cdots, n,$$

$$y_{jk} \in \mathrm{cl}S_j \cap (\bar{y}_j + \varepsilon_k \mathbb{B}), \quad j = 1, \cdots, m,$$

$$w_k = \sum_{i=1}^{n} x_{ik} - \sum_{j=1}^{m} y_{jk} \in \mathrm{cl}W \cap (\bar{w} + \varepsilon_k \mathbb{B})$$

和 $\varepsilon = \varepsilon_k, \widehat{N}$ 由 $\widehat{N}$ 替代时的对偶关系 (8.13)~(8.16). 很明显, 随着 $k \to \infty$, $(x_k, y_k, w_k) \to (\bar{x}, \bar{y}, \bar{w})$. 注意价格序列 $\{p_k^*\}$ 在 E^* 内有界. 由泛函分析的基本知识, 商品空间是 Banach 空间时可以得到该序列的弱 * 聚点 (在收敛子网的意义下) $p^* \in \mathrm{cl}^*\{p_k^* | k \in \mathbb{N}\}$. 如果 E^* 的闭单位球 $\mathbb{B}^*$ 是弱 * 列紧的 (比如 E 是 Asplund 或 β- 光滑 Banach 空间), 那么 $\{p_k^*\}$ 包括一个弱 * 收敛于某 $p^* \in E^*$ 的子序列. 对 (8.13)~(8.15) 取极限 (其中以 $\widehat{N}$ 取代 $\widehat{N}$), 则 p^* 在拓扑和序列两种情形皆满足极限关系 (8.46)~(8.48), 其中分别以由 $\widehat{N}$ 生成的拓扑和序列结构代替 $\overline{N}$.

剩下的是要证明, 如果集合 $\mathrm{cl}P_i(\bar{x})$, $\mathrm{cl}S_j$, $\mathrm{cl}W$ 中的一个在对应点是 $\widehat{N}$- 序列法紧的, 则可以选择 $p^* \neq 0$. 这对于具有弱 * 列紧对偶球的 Banach 空间是简单的. 但在一般 (非序列的) 情况下需要一些论证.

确切起见假设集合 $\mathrm{cl}W$ 在 $\bar{w}$ 是 $\widehat{N}$- 序列法紧的, 且 $p^* = 0$ 是 $\{p_k^*\}$ 的唯一弱 * 聚点. 那么随着 $k \to \infty$, 整个序列 $p_k^* \xrightarrow{w^*} 0$, 由对应于 $\widehat{N}$ 情形的 (8.15), 对所有 $k \in \mathbb{N}$, 存在某 $b_k^* \in \mathbb{B}^*$ 使得

$$p_k^* + \varepsilon_k b_k^* \in \widehat{N}(w_k; \mathrm{cl}W),$$

由此, 随着 $k \to \infty$, $p_k^* + \varepsilon_k b_k^* \xrightarrow{w^*} 0$. 由 $\mathrm{cl}W$ 的 $\widehat{N}$-SNC 性质, 这蕴涵 $\| p_k^* + \varepsilon_k b_k^* \| \to 0$, 从而 $\| p_k^* \| \to 0$. 这显然违背了定理 8.16 对于 p_k^* 的非平凡性条件 (8.16). 证毕.

△

很容易发现, 倘若 E 上闭凸子集的法锥结构与凸分析的法锥一致, 那么定理 8.8 的推论 8.9、推论 8.10 中的结果在定理 8.18 的抽象框架下都是成立的.

下面考虑具有有序商品空间 E 的经济. 首先看到, 在 (8.40)~(8.42) 给定的标准期望性或自由处置条件下, 倘若闭正锥 E_+ 是实的, 即是内部非空的, 则定理 8.18

所有的条件和 SNC 假设都是自动成立的. 这样就得到了下列有序商品空间情形的 Pareto 和弱 Pareto 最优配置的第二福利定理的抽象版本. 简洁起见, 这里只给出一般的 Banach 空间中拓扑法锥结构情形的结果; 若 $\mathbb{B}^* \subset E^*$ 是弱 * 列紧的, 其对应的序列情形的结果可类似给出, 如定理 8.18 的情形.

推论 8.19(有序商品空间中 Pareto 和弱 Pareto 最优配置的抽象第二福利定理) 设 E 为满足 $\operatorname{int} E \neq \emptyset$ 的有序 Banach 空间, 且定理 8.18 中拓扑法锥结构 $\overline{\mathcal{N}}$ 对 E 中的任意闭子集 Ω 满足 $\overline{\mathcal{N}}(\cdot;\Omega)$ 不大于 Ω 的 Clarke 法锥. 则下列断言成立:

(i) 给定 $\mathcal{E}$ 的局部弱 Pareto 最优配置 $(\bar{x},\bar{y})$, 假设以下条件之一成立, 即或者净需求约束集合 W 在 $\bar{w}$ 附近封闭且满足商品的自由处置条件 (8.40), 或者生产集合 S_j 中的一个在 $\bar{y}_j$ 附近封闭且遵从自由处置条件 (8.41). 那么存在一个满足关系 (8.46)~(8.48) 的非零边际价格 $p^* \in E^*$.

(ii) 给定 $n > 1$ 时的 $\mathcal{E}$ 的局部 Pareto 最优配置 $(\bar{x},\bar{y})$, 假设第 i 个消费者满足期望条件 (8.42). 那么存在一个满足 (8.46)~(8.48) 的非零边际价格 $p^* \in E^*$.

证明 容易看到, 对于任何 Banach 空间中的子集 Ω, 若对某非空开锥 K 成立包含关系 $\Omega + K \subset \Omega$, 则对每个 $\bar{x} \in \mathrm{cl}\Omega$, Ω 在该点附近具有上图 Lipschitz 性质. 这样, 当 $\operatorname{int} E_+ \neq \emptyset$ 时, 每个条件 (8.40)~(8.42) 都确保了相应集合的上图 Lipschitz 性质, 因此由命题 8.4 知定理 8.18 中的净需求 (对应地, 弱) 规范条件都成立. 由于在任何 Banach 空间中, 上图 Lipschitz 集合的 Clarke 法锥都是弱 * 局部紧的, 所以这样的集合具有关于这种锥的序列法紧性质. 因此只要所论的预法锥不大于 Clarke 法锥, 则 (8.40)~(8.42) 中对应的集合对这种预法锥就是序列法紧的. 这样一来所有定理 8.18 的假设都成立, 从而得到 Pareto 和弱 Pareto 最优配置的边际价格关系 (8.46)~(8.48). △

与 8.3.2 节中考虑的基本情形相似, 下面的结论指出, 强 Pareto 最优配置第二福利定理的确切版本不需要净需求约束条件; 对凸或非凸的经济学, 即使其有序商品空间的闭正锥的内部是空的, 该定理也可以成立.

定理 8.20(强 Pareto 最优配置的确切第二福利定理的抽象版本) 设 $(\bar{x},\bar{y})$ 为具有有序 Banach 商品空间 E 的经济 $\mathcal{E}$ 的局部强 Pareto 最优配置, 且集合 S_j 和 W 分别在 $\bar{y}_j$ 和 $\bar{w}$ 附近局部闭. 假设下列集合中的一个:

$$\mathrm{cl}P_i(\bar{x}),\quad i=1,\cdots,n,\quad S_j,\ j=1,\cdots,m,\quad W$$

在相应点上是 $\widehat{\mathcal{N}}$-SNC 的, 其中抽象预法锥结构 $\widehat{\mathcal{N}}$ 满足假设 (H1) 和 (H2), 并且下述条件之一成立:

(a)E_+ 是生成的, 并且自由处置/期望性条件 (8.40)~(8.42) 中的一个得到满足;

(b)$\operatorname{int} E_+ \neq \emptyset$, 集合 W, S_j, $j=1,\cdots,m$, 和 $P_i(\bar{x})$, $i=1,\cdots,n$ 中至少两个满足 (8.40)~(8.42) 的相应条件.

那么存在一个对偶元素 $p^* \in E^* \setminus \{0\}$, 满足边际价格关系 (8.46)~(8.48), 其中 $\overline{\mathcal{N}}$ 表示 $\widehat{\mathcal{N}}$ 生成的拓扑法锥结构. 此外, 如果 E^* 的闭单位球 $\mathbb{B}^*$ 是弱 * 列紧的, 则 (8.46)~(8.48) 中的拓扑结构 $\overline{\mathcal{N}}$ 可以由 $\widehat{\mathcal{N}}$ 生成的序列法锥结构 $\mathcal{N}$ 替代.

证明　类似定理 8.18 的证明, 对定理 8.17 中关于强 Pareto 最优配置的近似关系取极限即可. △

对 2.5.3 小节讨论的某些 Banach 空间上的特殊的预法锥和法锥结构, 本节建立的抽象结果都可以有效地具体化.

8.4.2　公共商品及交换限制

结尾这一小节简要探讨本章的方法和结果在具有公共商品的经济学上的扩展, 也提到一些该方法在具有公共环境和交换限制的竞争均衡模型上的一些可能的应用.

与上述研究的福利经济模型不同, 公共商品经济包括两大类括商品: 私有的和公共的. 第一种商品的消耗是排他性的, 即任何商品一旦为某消费者占有, 则自动地不再可为其他人获得. 相反, 如果消费对任何人都是均等的, 那么商品就是公共的. 数学上这意味着商品空间 E 可表示为两个 Banach 空间的乘积 $E = X \times Z$, 其中 X 和 Z 分别是私人和公共商品的空间. 这样消费变量 $x_i \in X$, $i = 1, \cdots, n$ 表示私人商品, 而变量 $z_i \in Z$, $i = 1, \cdots, n$ 与公共商品对应; 如上, $y_j \in S_j \subset E$ 表示生产变量. 简洁起见, 考虑 (8.1) 中“市场出清”的情形, 其中稀有资源 $\omega \in X$ 的初始禀赋仅对私有商品给出, 此时可把涉及私人和公共商品经济的市场约束陈述为

$$\sum_{i=1}^{n}(x_i, z_i) - \sum_{j=1}^{m} y_j = (\omega, 0) \tag{8.49}$$

注意, 在约束条件 (8.49) 中没有公共商品的禀赋, 这是公共商品经济最关键的特点.

沿着上面的没有公共商品的经济的研究思路, 把市场限制条件 (8.49) 结合到 (8.18) 中集合 Ω_2 的结构中, 并运用极点原理, 即可得公共商品经济学的相似结果. 这些结果包括了 Pareto 最优配置的三种类型的扩展第二福利定理的近似和确切版本, 也包括了 8.4.1 节提出的这些定理的抽象版本. 与本章的基本结果相比, 公共商品经济的主要改动在于 (这里只给出定理 8.8 中确切/极限条件的情形): 与存在满足 (8.30) 和 (8.31) 的非零边际价格 $p^* \in E^*$ 不同, 这里存在价格 $p^* = (p_x^*, p_z^*) \in X^* \times Z^*$, $p_i^* \in Z^*$ $(i = 1, \cdots, n)$, 其中 $(p_x^*, p_i^*) \neq 0$ 至少对一个 $i \in \{1, \cdots, n\}$ 成立, 且

$$-(p_x^*, p_i^*) \in N\big(\bar{x}_i; \mathrm{cl}P_i(\bar{x})\big), \quad i = 1, \cdots, n, \tag{8.50}$$

$$(p_x^*, p_z^*) \in N(\bar{y}_j; \mathrm{cl}S_j), \quad j = 1, \cdots, m, \tag{8.51}$$

$$p_z^* = \sum_{i=1}^{n} p_i^*. \tag{8.52}$$

可以看出, 条件 (8.50) 和 (8.51) 只是在商品空间 $E = X \times Z$ 时 (8.30) 和 (8.31) 的具体化. 而最后一个条件 (8.52) 则证实了可以追溯到 Samuelson[1189] 的具有公共商品空间的福利经济学的基本结论: 在 Pareto 最优配置中, 公共商品的边际转换率等于个人边际替代率之和.

如上文所述, 本章研究福利经济学模型的 Pareto 优化理论的方法的主要数学工具是变分分析的极点原理, 所涉及的集合系统在没有公共商品的模型中由 (8.17) 和 (8.18) 定义, 而对公共商品经济则按 (8.49) 修改而来. 这样的考虑也适用于具有所谓公共环境的福利模型, 这涉及私有市场和非市场 (如法律系统) 部门的结合; 更多的细节和例子见 Villar[1287].

从数学上来讲, 这样的模式可用类似 8.4.1 节的具有依赖于参数的消费、生产、偏好的基本福利模型来描述. 在 Asplund 商品空间的情形, 这种模型的适当的扩展第二福利定理可见 Habte 的博士论文 [533], 该研究也是基于极点原理的.

可以看到, 上述考虑的所有模型不涉及中介间交换的任何限制. 这样的限制曾出现在其他一些竞争性经济均衡的模型内, 例如 Makarov, Levin 和 Rubinov[829], Rubinov[1182]. 从数学上来讲, 这些模型的大部分可以用类似上面形式来描述, 只是与 (8.1) 和 (8.49) 相比市场约束条件中的消费和生产变量之间的关系更加复杂. 这导致了在相应的极值系统中, 市场制约集合 Ω_2 要做相应的改动, 然后就可由变分分析的极点原理来处理. 该类型的一个重要的直接分配模型见 Habte[533], 该文导出了该模型广义第二福利定理的各种版本.

8.5 第 8 章的评注

8.5.1 福利经济中的竞争均衡和 Pareto 最优

竞争均衡模型和消费者和生产者/企业市场价格分散的基本思想可以溯源到 Léon Walras, 在他 1874~1877 年出版的开创性著作《纯粹政治经济学要义》[1300] 中, 回答并证明了一些他的前辈们提出的一些重大问题. 特别地, 这包括 Adam Smith[1217] 提出的"为什么大量出于自身利益而作出独立决策的经济行为者在私有制经济中不会造成社会混乱"的问题. Smith 本人对于消费者和企业的市场行为的客观协调方面有深刻的见解, 并得出"看不见的手"这一著名的结论. 然而, 只有数学模型才可以提供实证观察和结论的科学论证. 通过建立这样一个模型, Walras 为竞争经济的一般均衡理论奠定了基础, 而这种经济亦称为福利经济.

大致说来, 竞争价格均衡是这样一种情形, 此时所有经济行为者在给定的价格

下同时实现他们的计划. 从 Pareto 的著作 [1053] 开始, 人们早就认识到, 福利经济学的竞争均衡和可行的资源配置效率的适当的概念之间有一定的联系; 经济学中效率的概念一般即为 Pareto 最优性. 根据经典的 Pareto 原理, 一个可行的配置比另一个更优越, 如果它是所有经济行为者的首选; 也就是说"优越"指的是一致的协议. 关于 Pareto 最优性在经济模型有用有几个不同的版本和现代诠释; 特别参看定义 8.2 和下面的评注.

可能首次严格论证福利经济学模型中 Pareto 最优的必要条件的是 Lange[738]. 他建立了在任何 Pareto 最优配置中消费与生产部门的边际替代率之间的相等关系; 对以往尝试的评论、相关结果, 和未发表的材料亦可见 Hicks[566], Samuelson[1188], Khan[671]. Lange 的证明是基于这样的观察, 即在特定的规范条件下, Pareto 最优的资源配置可以被解释为某些非线性规划约束问题的最优解, 这允许他使用经典的 Lagrange 乘子法则, 当然他考虑的是有限维商品空间, 并且假设了效用和生产函数的标准光滑性质. 这一结果现在被称为福利经济学的第二基本定理的原始版本, 该名字首见于后来的凸经济学中的 Arrow-Debreu 模型中. 事实上, 在可微性和凸性假设下, Lange 的结果可由凸经济的 Arrow-Debreu 模型中 Pareto 最优的必要条件导出, 那里这称为"第二福利定理", 具体见下文.

8.5.2　福利经济学的凸模型

初创在 1951 的论文 Arrow[26] 和 Debreu[309] 中的所谓的 Arrow-Debreu 一般均衡模型不仅是数理经济学及其应用的基石, 也极大地影响了数学中的优化相关领域, 特别是凸分析. 关于有限维与无限维商品空间的凸经济的 Arrow-Debreu 模型的各个方面的研究文献很多, 例如参见 Aliprantis, Brown, Burkinshaw[10]、Debreu [310]、Florenzano[459], Mas-Colell, Whinston, Green[856] 及其所录文献.

在他们的模型中, Arrow 和 Debreu 把焦点从替代边际率转移到作为价格均衡的 Pareto 最优配置的分散化. 在凸性和相关假设下, 由 Pareto 最优与其竞争价格的等价性, 他们建立了 Pareto 最优必要和充分条件, 其中的竞争价格均衡理解为这样的价格, 它使消费者支出最小化和生产商利润最大化刚好在给定的 Pareto 最优配置处实现. 此外, 这样的价格均衡的存在性是在恰当的假设下对一般的凸模型证明的, 其中的商品空间可以是无限或有限维的. 证明这些结果的主要数学工具是导出最优性条件的分离定理和建立均衡存在的不动点定理 (Brouwer 和 Kakutani). 这些定理都基于凸性.

第一、二福利定理的名字是从 Pareto 最优化和价格均衡之间的等价性的相应部分中产生的. 第一福利定理指出, 任何均衡配置是 Pareto 最优的 (Pareto 最优的充分条件); 第二福利定理指出了其反面, 即任何 Pareto 最优配置提供了一个价格均衡 (Pareto 最优的必要条件). 可以看到, 第一福利定理极大地依赖于凸性; 在非

凸 (甚至光滑) 模型中不具有任何类似结果. 但是, 第二福利定理允许在非凸模型上的深远扩展; 这实际上是第 8 章主要内容. 还要注意, Arrow-Debreu 模型是基于凸性, 不需要任何像以前研究中的可微性假设.

除了上述凸性假设, 具有无限商品的经济的 Arrow-Debreu 模型需要所涉及的某些集合的内部非空条件. 从数学上来讲, 这是由于在无穷维空间应用分离定理的缘故. 在有序的拓扑空间的情况下, 内部非空条件化归为所论商品空间的正锥/正象限的内部非空假设, 但这在许多重要的经济建模情形下是不成立的. 为避免这种限制, Mas-Colell [855] 提出了著名的凸经济适宜性假设, 其有序的商品空间是可能无内点的拓扑向量格. 对于具有有限维与无限维商品空间的凸经济 Mas-Colell 的适宜条件的各种扩展和改进可以在众多论著中找到, 如 Aliprantis、Tourky 与 Yannelis[14], Florenzano[459], Mas-Colell, Whinston 与 Green[856], Tourky[1261] 及其所录文献. 这里最后提一下最近 Naniewicz 的论文 [993], 其在自反商品空间上给出了发展 Arrow-Debreu 模型一个全新方法, 其中只需要通常的凸性而不需要内点条件. 这基于把经济模型归结为一个变分不等式系统然后运用伪单调多值映射的理论.

8.5.3 进入非凸领域

如 8.1.1 小节已经提到的, 凸性的相关性假设在许多重要的经济应用中是否有意义是值得怀疑的, 这在 Arrow-Debreu 模型形成之前就已经被认识到了; 见上文提到的出自 Samuelson[1188] 的引述. 事实上, 各种类型的非凸性不可避免地出现在垄断竞争, 规模效益增长、不完全市场和外部因素等等的建模中. 更多的例子和讨论参看 Anderson[18], Cornet [287], Florenzano[459], Khan [671], Quinzii [1113], Villar [1287] 及其所录文献.

在他的福利经济学非凸模型的价格分散化的开创性研究中, Guesnerie[524] 以非凸经济学 Pareto 最优配置的必要最优条件的形式建立了广义版本的第二福利定理. 他没有假定初始生产和偏好集合的凸性, 而是假设了其 (局部) 切向近似的凸性, 然后使用经典的凸锥分离定理. 具体的过程中他利用了一般优化理论中 Dubovitskii 和 Milyutin[370] 所提出的“内部位移锥”这一概念.

Guesnerie 的研究非凸经济学 Pareto 最优性的方法在涉及有限维与无限维商品空间的非凸经济学中有很多扩展, 例如参看 Bonnisseau 和 Cornet [135], Brown [181], Cornet [286], Khan,Vohra[673, 674], Quinzii[1113], Villar[1287] 及其所录文献. 这些研究大多都使用 Clarke 切锥, 好处是这种锥自动就是凸的, 因此可以采用经典的凸分离定理. 此时边际价格 (对应于非光滑和非凸模型的边际替代/转换率) 通过对偶的 Clarke 法锥表示. 然而人们后来认识到, Clarke 法锥经常太大而不足以描述边际价格; 见 Jouini [642] 和 Khan[671] 中的例子和讨论.

在论文 [671] (它的第一个版本是 1987 年的预印本) 中, Khan 对具有有限维商品空间的非凸经济得出了令人满意得多的第二福利定理的扩展, 其中的边际价格由本书中的基本法锥表示. 要注意到他的方法不涉及任何凸分离, 而是归结为 Mordukhovich[892] 建立的非光滑数学规划必要最优性条件 (Lagrange 乘子); 请比照 5.1.3 节的定理 5.21(iii) 在有限维的情形. 在这方面, Khan 的方法标志着由 Arrow 和 Debreu 首创的分离方法回归到了 Hicks, Lange, Samuelson 在福利经济学基础中和所倡导的经典的"Lagrange 乘子"的思想. 有限维中的由基本法锥描述的广义第二福利定理一个类似的版本是由 Cornet [288] 在一种不同的经济模型中获得的. 对相应的最大化问题的必要最优条件, 他给出了一个直接的证明.

对具有有限维或无限维商品空间的福利经济学的各种非凸模型的第二福利定理进一步的研究包括 Borwein, Zhu [164], Flåm [452], Flåm 与 Jourani [453], Florenzano, Gourdel, Jofré[460], Habte [533], Jofré [633], Jofré, Rivera [635], Khan [669, 670], Malcolm, Mordukhovich [836], Mordukhovich [920, 922, 930], Villar [1288], Zhu [1375], 他们使用了基本法锥及其无限维推广、修改、和抽象版本. 下面将详细讨论.

8.5.4　极点原理和福利经济学模型非凸分离

Mordukhovich [920, 922] 提出一种研究非凸经济 Pareto 最优和导出第二福利定理的新方法, 即利用第 2 章中完整讨论的变分分析极点原理. 如前述, 极点原理为闭集系统的局部极值点提供了最优性必要条件, 这种极点特别包括了约束数学规划和矢量优化问题局部解. 另一方面, 它给出了非凸集情形下经典凸分离的一个变分对应, 其于非凸变分分析的角色正如凸分离定理之于凸分析. 因此, 基于极点原理 Pareto 最优性的方法可以被视为前面讨论的光滑模型和凸模型两个方法的统一. 可以看到, 第 8 章所有的结论都基于极点原理.

在福利经济学模型的 Pareto 最优的研究中, 与上述密切相关但略有不同的方法是由 Jofré [633] 提出的. 该方法基于 Borwein, Jofré [148] 建立的集合的和的边界点的次微分条件. 此集合和的特性在文献 [148] 被作为一个非凸分离, 实际上相当于另一集合系统的局部极值性质. 进一步, 在文献 [148] 所获得的集合和的边界点的次微分特征事实上与极点原理近似版本等价; 更精确陈述和讨论见近期的文章 Kruger [716] 和 Zhu[1375]. 前述的文献 [164, 452, 453, 460, 533, 633, 635, 836, 920, 922, 930, 1375] 所获得的非凸经济扩展的第二福利定理的结果, 或是通过直接应用极点原理, 或是通过使用等价的 Borwein, Jofré [148] 建立的边界/非凸分离性质而得到.

8.5.5　基本模型及解的概念

在 8.1.1 小节描述的福利经济学一般模型的 $W = \{\omega\}$ 的情形在当代微观经济学文献受到广泛认可, 其中 ω 是已知的稀缺资源的初始总禀赋, 例如参看书籍

Aliprantis、Brown 与 Burkinshaw[10], Mas-Colell,Whinston 与 Green [856]. 请注意, 经济 $\mathcal{E}$ 的偏好关系是由集值映射 P_i 所给出来的, 没有使用预先定序、效用函数及其他经典福利经济学的常规属性; 请比照 Debreu[310].

当 $W = \{\omega\}$, 可行配置定义 8.1 中的关键关系 (8.1) 归结为所谓的市场出清条件. 如 Mordukhovich [920, 922]、Malcolm 与 Mordukhovich [836], 引入净需求约束集 W 的好处至少有以下两方面:

— 它允许同时考虑标准情形 $W = \{\omega\}$ 和通过有序商品空间 E 的闭正锥 $E_+ \subset E$ 定义的所谓的商品的隐含自由处置 $W = \omega - E_+$(参看 Cornet [288]).

— 由于 W 是任意的, 可行条件 (8.1) 能够反映出初始总禀赋不明朗时 (比如因为信息不全) 市场的不确定性; 请特别比照 7.5.19 节中不确定条件下极大极小控制问题的讨论.

定义 8.2(i, ii) 中的 (局部) Pareto 和弱 Pareto 最优配置概念在经济学文献中是标准的. 它们与一般向量/多目标优化问题中的 Pareto 和弱 Pareto 最优性的传统概念是一致的; 例如参看 5.3 节、其评注, 及相应的文献.

据我们所知, 定义 8.2(iii) 中所给的福利经济学模型的强 Pareto 最优配置概念是由 Khan [670] 首先明确提出并探讨的. 根据与 Ali Khan 的私人交流, 这与 Debreu 的工作 [311] 一致, 其在 Khan 和 Rashid [672] 的渐近性研究中已露出端倪, 而该文的源动力则来自 Hildenbrand [568]. 对文献 [672] 中方向的进一步研究见 Anderson [18].

8.5.6 规范条件

如上所述, 特定的约束规范条件不仅出现在所有有关 Pareto 和弱 Pareto 最优配置的深入研究中, 而且出现在福利经济学光滑和凸模型的第二福利定理的最初版本中. 强赋在 Arrow-Debreu 凸模型 (有几种版本和修改) 的这种类型的关键条件作为 (非空) 内部条件为人们所熟知. 定义 8.3 中的净需求规范条件可被视为福利经济学非凸模型的传统内部条件的深远扩展.

分别加在 Pareto 和弱 Pareto 最优配置上的两个条件 NDQ、NDWQ 都首见于 Mordukhovich [920] 中, 而以前的 $W = \{\omega\}$ 及生产集 S_j 是闭集情形的 NDQ 条件是 Jofré 在文献 [633] 给出的 (还见于以前其与 Rivera 合作的预印本中; 见文献 [633] 和 [635] 及其所录文献), 名为"渐近包括条件". 此外, 如果此时序列 e_k 被 αe 取代, 其中 $e \in E$ 是固定向量, 而 $\alpha > 0$ 足够小, 则 NDQ 条件 (8.3) 归结为可追溯到 Cornet [286] 的规范条件, 称为"径向 Lipschitz 条件"(如 Bonnisseau 和 Cornet [135]) 或"Cornet 约束规范"(如 Khan [671]); 这些论文中还可找到更多参考资料和讨论. 关于弱 Pareto 最优配置的 NDWQ 条件的进一步研究和应用, 请读者参看最近工作 Zhu [1375]、Borwein 与 Zhu [164].

命题 8.4 的完整扩展版本由 Malcolm 和 Mordukhovich [836] 建立, 平行地给出了 NDQ 和 NDWQ 性质的易验证的充分条件. 而在 $W = \{\omega\}$ 情形下的一些以前的版本, 则见于 Bonnisseau 和 Cornet [135]、Cornet [286]、Jofré [633] 以及 Khan [670, 671]. 值得注意的是, 命题 8.4 的断言 (i) 是 Mas-Colell[854] 中期望方向条件的直接推广, 它相关于凸经济学中经典的“越多越好”条件, 其中的商品空间由其具有非空内部的闭正锥定序.

正如定义 8.3 之后讨论的那样, 如果偏好与生产集中至少一个是局部上图 Lipschitz 的, 则 NDQ 和 NDWQ 两个条件都是自动成立的. 对于凸集, 这等价于经典的内部非空条件. 集合的上图 Lipschitz 性质被 Khan [671] 称为“在某些方向饱满”性质. 另外也注意, 推出 NDQ 和 NDWQ 条件里集合和中的项趋于提高该集合和的上图 Lipschitz 性质, 特别是在市场经济行为者数量较大的情况下.

由命题 8.4, 本章的模型中使用净需求约束集合 W 的好处之一是, 即使要求 W 是上图 Lipschitz 的, 也可在偏好和生产集上避免任何额外要求 (这在 $W = \{\omega\}$ 时是没问题的). 由此就自动包含了有限维商品空间情形下 Cornet [288] 研究的涉及所谓的“自由处置 Pareto 最优”的福利模型.

可以看到, 本章中的规范条件与 8.5.2 小节探讨的带有有序商品空间的凸模型的 Mas-Colell 适宜条件及修改无关. Mas-Colell 适宜性的一些非凸版本, 最近由 Florenzano、Gourdel 和 Jofré [460] 在有序商品空间被赋予了 Banach 格结构的模型中, 在广义第二福利定理的弱 Pareto 最优框架下, 引入并研究. 但本章的研究从未用过格结构.

8.5.7　第二福利定理的近似版本

为避免上述的非空内部与适宜性假设, 在凸性假设下处理 Pareto 和弱 Pareto 最优配置的各种微观经济模型中, 曾有几个第二福利定理的近似版本及其经济解释; 请特别参考 Aliprantis 与 Burkinshaw [11], Hildenbrand [568], Khan 与 Rashid[672], Khan 与 Vohra [675]. 其中文献 [675] 用到了 Banach 空间中关于凸集支撑点在边界上稠密的著名的 Bishop-Phelps 定理 [116], 这一定理被 Ekeland [399] 视为现代变分原理的始祖.

根据 Mordukhovich [920, 922], Malcolm 与 Mordukhovich [836], 8.2.1 小节中对非凸经济学的 Pareto 和弱 Pareto 最优配置建立了第二福利定理的近似版本, 其中在 Asplund 空间的情形, 边际价格如文献 [836, 922] 由 Fréchet 法锥表示, 而在一般恰当的 Banach 空间中, 则如文献 [920] 由更一般的预法锥结构表示 (亦见 8.4.1). 这些结果的证明基于近似极点原理的相应版本, 它可视为 Bishop-Phelps 定理的非凸扩展; 参看命题 2.6, 推论 2.21 和 2.6.4 小节的后续讨论.

Jofré 在文献 [633] 中也建立了该类型的结果, 即 Banach 空间上非凸经济 Pareto

最优配置的第二福利定理的“黏性”版本. 对“市场出清”的福利模型, 基于“渐近包括”规范条件, Jofré 通过由公理描述的、满足一般要求的次微分形式, 建立了近似/黏性第二福利定理的次微分形式. 然而, 并非所有这些要求对于 Banach 空间中 Fréchet 次微分都是满足的, 这与 Mordukhovich [920] 是不同的. 所以 8.2.1 小节提出的结果不能从文献 [633, 定理 2] 获得, 因为该定理基于 Borwein 和 Jofré [148] 的非凸边界条件; 参看 8.5.4 小节.

8.5.8 法紧条件下第二福利定理的确切版本

第二福利定理的确切 (或点基) 版本, 是指类 Pareto 最优必要条件中的边际价格是以最优点处 (而不是附近) 的法锥结构表示的. 这种类型的结果对经济应用是最需求的; 它们包括光滑或凸假设下的有限与无限维商品空间上各类经济学的第二福利定理的所有经典版本. 对非凸经济学, 绝大多数“确切”结果的边际定价规则是由 Clarke 法锥描述的. 用到基本法向量的改进版本见 Khan [670] 和 Cornet [288], 其中的商品空间是有限维的; 更多的细节和讨论见 8.5.3 小节.

众所周知, 对带有无限维商品空间的福利经济学, 第二福利定理的确切版本需要额外的紧性条件, 这和无限维空间中一般优化理论完全一致; 请比照第 5 章. 可以看到, 对于凸经济学而言, 传统的内部非空与适宜性的假设已包含了足够的紧性.

然而非凸经济学的情形是不一样的, 大多数确切或基于点类型的结果都明确假设模型中涉及的某些集合的上图 Lipschitz 性质; 特别可见 Bonnisseau 与 Cornet [135]、Khan [670]、Khan 与 Vohra [673, 674]. 据我们所知, 该性质是经典的内点条件在非凸情形的恰当的扩展.

8.2.2 和 8.3.1 小节中的确切类型的结果见于 Mordukhovich [922]、Malcolm 与 Mordukhovich [836], 而 8.4.1 小节中给出的 (不完整) 抽象本取自 Mordukhovich[920]. 8.2.2 和 8.3.1 小节中的第二福利定理的扩展版本通过 Asplund 商品空间中的基本法锥来描述边际定价. 在 8.4.1 小节里恰当的 Banach 空间情形中, 该法锥可以被其抽象的拓扑或序列极限的抽象版本代替.

8.2 至 8.4 节中探讨的所有扩展型第二福利定理确切版本的一个显著特征是, 对所用的基本序列法紧 (SNC) 性质或其抽象的序列版本, 只需要求对所考虑的福利模型的偏好、生产或净需求约束集中的一个成立即可. 这与书中研究的其他无限维中涉及有限多集合的确切/极限结果的类似情形有显著的差异, 这其中包括确切极点原理、广义微分法则、约束优化的必要条件. 事实上, 在先前的所有情形下, 都要求除一个外, 所论情形中所有其他集合都具有 SNC 和/或相关性质. 之所以能在所讨论的经济模型中取得如此重大的改进, 是因为应用极点原理的极值系统中的约束集合 (8.18) 具有特定的线性分离结构.

对在这个方向上的结果, 这里提一下 Jofré[633] 以及有类似结果的 Flåm 和

Jourani[453], 他们通过在适当的 Banach 空间上的距离函数的抽象极限次微分, 得到了第二福利定理的确切次微分版本. 这些结果假设了福利模型中的一个集合的紧上图 Lipschitz 性质. 从 1.1.4 小节可知, 紧上图 Lipschitz 性质碰巧是 SNC 性质的一个拓扑对应. 但是不仅在没有任何分离性结构的一般 Banach 空间上紧上图 Lipschitz 性质严格强于 SNC 性质, 而且在闭对偶球是弱 * 列紧的空间里亦如此, 这特别包括了 Asplund 空间 (其中如此的集合甚至可以是凸集, 比如 3.1.1 小节中范例 3.6 中的集合). 读者在 Flåm [452]、Florenzano, Gourdel 和 Jofré [460] 中可以找到紧上图 Lipschitz 假设下确切第二福利定理的更多扩展.

8.5.9　有序商品空间中的 Pareto 最优性

考虑具有有序商品空间的非凸经济学, 一个自然的问题就是满足确切扩展第二福利定理的关系的边际价格 p^* 是否具有正性, 即 $0 \neq p^* \in E^*_+$, 其中 $E^*_+ \subset E^*$ 是为经济 $\mathcal{E}$ 的商品空间 E 定序的对偶正锥, E^* 是 E 的对偶空间. Malcolm 与 Mordukhovich [836] 发现, 如果对应的偏好, 生产, 或净需求约束集合满足定理 8.13 中的传统的期望、自由处置、或隐含自由处置假设中的一个, 则所需的价格正性可由定理 8.8 中的关于边际价格 p^* 的任一个 "确切" 法锥关系 (8.30)~(8.32) 得出. 这是文献 [836] 引理 8.12 给出的 "自由处置" 集合的法锥的正性的直接推论.

应该再次强调, 以上讨论的涉及经济学中 (局部)Pareto 或弱 Pareto 最优配置第二福利定理所有结果, 都依赖于对应的约束规范条件, 即 NDQ 和 NDWQ 条件, 这把经典的内点条件扩展到非凸经济. 因为任何强 Pareto 最优配置显然也是 Pareto 最优, 所以所获得的结果对具有更多限制的强 Pareto 最优亦成立.

如上所述, 福利经济学模型中的强 Pareto 最优概念是 Khan [670] 提出的, 他利用此概念推导出第二福利定理确切版本, 其中边际定价由 Ioffe [597] 在局部凸线性拓扑空间中定义的 "近似" 法锥 (A- 法锥) 来描述; 该法锥是本书中有限维中基本法锥在无限维的一个推广. Khan [670] 的主要结果证明的关于非凸经济强 (局部)Pareto 最优的扩展第二福利定理中, 有序的商品空间假定为格, 具有自反的偏好关系, 且 "自由处置" 净需求约束集具有形式 $W = \omega - E_+$. 此外, 文献 [670] 中还假定, 对于所有的 $i = 1, \cdots, n$, 所有 $j = 1, \cdots, m$, 定理 8.13(a) 和 (b) 的期望和自由处置条件都满足, 并且在所论的强 Pareto 最优配置附近每一个生产、偏好集合都具有上图 Lipschitz 性质.

对具有任意有序 Banach 商品空间的非凸经济, 由定理 8.13(Asplund 情形) 和推论 8.19(抽象情形) 可知, Khan 的第二福利定理的改进版本, 在显著减少的假设下, 对于任意弱 Pareto 最优配置成立, 并不只是针对强 Pareto 最优配置. 事实上, 综合命题 8.4(ii) 建立的 NDWQ 性质的充分条件可得, 在 Khan 理论框架中对于任何弱 Pareto 最优配置扩展第二福利定理都成立 (其中只需 Banach 空间结构, 不需

格结构), 如果偏好、生产集合中只有一个在相应点周围是上图 Lipschitz 的, 而同时定理 8.13 的期望条件 (a) 和自由处置条件 (b) 两者都不必满足. 此外, 在这些较弱的假设下, 本节上述结果改进了 Khan 的边际定价规则, 因为此时在 Asplund 情形可用基本法锥、或在一般 Banach 空间上用 Ioffe 的 G- 法锥, 两者都小于 Khan 使用的 Ioffe 的 A- 法锥.

在 Khan 的另一篇文章 [669] 中, 他 (基于 Treiman [1262] 中的一些结构) 构造了一个具有传统 Asplund 商品空间 c_0(收敛到零的实数序列空间, 赋予上确界范数) 的非凸经济例子, 其中生产集合在 Pareto 最优配置的 Ioffe 的 A- 法锥是整个对偶空间 ℓ^1. 这个例子另一个显著特点是, 不仅本书中的基本法锥, 而且其弱 $*$ 凸化 (即 Clarke 法锥) 在该例中都严格地小于 Ioffe 的 A- 法锥, 因此可提供广义第二福利定理框架中非平凡的边际价格信息. 注意文献 [670] 中 Khan 的结果不适用该例, 因为生产集合不是上图 Lipschitz 的, 不呈现商品的自由处置.

8.5.10 没有规范条件的强 Pareto 最优性

出人意料的是, 强 Pareto 最优配置在有序商品空间福利经济模型 (包括凸和非凸) 中扮演了一个特殊的角色, 即对于第二福利定理近似与确切版本, 它们无需任何净需求规范条件 (包括非空内部与适宜条件). 这是 Mordukhovich[920, 922] 观察到的, 然后在最近的文献 [930] 得到发展; 在 Asplund 空间情形, 对应的材料在 8.3.2 小节中由本书的基本结构给出, 而在一般 Banach 空间上抽象法/次微分结构的总体框架下, 相应的结果放在 8.4.1 小节中. 对注 8.15 讨论的改良的强 Pareto 最优配置, 它基于与 Glenn Malcolm 的个人交流, 写作本书时这些材料尚未发表.

这里首先强调, 针对 Pareto 和弱 Pareto 最优配置在 8.2 节中建立的第二福利定理的方法中, 规范条件只是为了使这样的配置可以归结为一些集合系统的局部极值点, 然后就可以利用极点原理和适当的分析法则. 然而在类自由处置条件下、具有有序商品空间经济中的强 Pareto 最优配置的情形中, 这个简约过程并不需要这样的约束规范条件 (这种情形与正锥的非空内部紧密相关), 这是因为强 Pareto 最优性本质上是不同的, 它直接蕴涵着相关集合系统的极值性质.

更确切地说, 完成这样的约简需要在生产、偏好和净需求约束集中至少两个满足上述自由处置/期望假设. 否则, 有序空间的正锥在 $E = E_+ - E_+$ 意义上需要是生成的, 这似乎不是一个很大的限制; 例如, 在任何 Banach 格 (或 Riesz 空间) 它都成立. 值得注意的是, 在生成正锥情形, 强 Pareto 最优配置归结到局部极值点的简约过程是基于有序 Banach 空间理论中高深的 Krein-Šmulian 定理.

8.5.11 非线性定价

众所周知, 通常用于经济建模并且参与第二福利定理多种版本的影子价格在数

学上被解释为商品空间上的对偶向量或线性连续泛函. 从这一角度, 它们被称为线性价格.

为实现市场均衡和其他福利经济学需要的基于价值的特征, 在经济学文献中, 尤其是关于价格歧视、累进关税、土地市场及组合交易的模型中, 已经探索了非线性价格的使用. Arrow 与 Hurwicz [27] 可能是关于非线性定价的最早文章之一.

最近, Aliprantis, Tourky 与 Yannelis[14] 为不具有商品空间的格结构的福利经济学凸模型初创了非线性定价的一种新方法. 其最大的动力来自于, Mas-Colell 没有内点假设的福利经济学的基本理论严重地依赖于商品空间的格性质, 甚至在有限维情形也是如此. 这里特别向读者推荐 Aliprantis、Monteiro 和 Tourky 的论文 [13], 它包含一个有两个交易者、和没有格结构三维商品空间的凸经济的一个引人注目的例子, 里面没有 Walras 均衡, 并且第二福利定理不成立.

正如 Aliprantis、Tourky 与 Yannelis[14] 所指出的, 与以前研究中应用的线性价格相比, 新的非线性价格理论为不具有商品格结构的有限维和无限维凸模型提供了一个足够的一般均衡理论. 注意, 文献 [14] 中所用的非线性价格尽管可能是非光滑的, 但总是凹的和正齐次的. 此外, 在向量格中它们归结为标准线性价格; 更多的细节和参考资料见文献 [14], 进一步发展和应用见 Aliprantis、Tourky 与 Florenzano[12].

8.2 节提出了非线性定价完全不同类型的一些结果 (特别参看定理 8.7 和注 8.11); 读者可以在 Mordukhovich [930] 发现更多的结果和讨论; 而对边际定价的这样一个非线性价格的解释可能首见于 Malcolm 与 Mordukhovich [836, 推论 4.3]. 该方法不涉及商品空间的格甚至序结构, 但显示了凸模型和非凸模型下第二福利定理的差别.

事实上, 凸经济学的线性价格支持一个完全分散均衡, 即在 Pareto 最优配置中, 每个公司实现利润最大化、每一位消费者达到支出最小化. 第二福利定理非线性边际定价版本只能提供在最优配置点由法锥给出的线性价格的局部描述. 然而, 如果使用“光滑非线性价格”, 则可以支持完全非凸模型中对所有偏好和生产集的全局分散 (凸类型, 或最大化–最小化) 均衡. 此外, Pareto 最优配置中, 这些非线性价格变化率 (即导数) 任意地接近来自于近似第二福利定理的线性边际价格.

数学上, 这些非线性价格的结论是基于定理 1.30 中获得的 Fréchet 法向量的光滑变分描述. 该定理包含变分分析中很强的几何结果, 其证明涉及变分原理. 必须强调, Fréchet 法向量结构对于这样的变分描述至关重要; 这些结果, 与 8.2 节和 8.3 节中扩展第二福利定理的其他近似与确切版本不同, 没有 8.4 节提出和讨论的抽象类似结果.

另一方面, 上述的 Fréchet 法向量的变分描述使得可以从 8.2 节和 8.3 节中获得的广义第二福利定理的确切版本中, 得出边际价格的有用的、有经济价值的诠

释. 由于在 Asplund 空间中基本法向量可由 Fréchet 法向量逼近, 定理 8.8 中建立的、8.3 节中改进确切版本中也使用的边际价格均衡关系 (8.30)~(8.32) 可以被解释为通过非线性价格实现的非凸模型中的极限分散均衡. 这里可以看到这种基于非线性价格的极限分散均衡与最近由 Jofré、Rockafellar 与 Wets [636] 对交换的凸 Walras 模型提出的所谓的"虚拟均衡"的一些相似性, 后者是对基于线性价格的经典均衡取极限得到的. 他们的方法是基于约简为非单调变分不等式. 这些结果在随后的论文 [637] 中进一步扩展到带有市场交易的、消费和生产的更一般性的凸 Walras 模型.

可以看到, 在以前的福利经济非凸模型中若以 Clarke 法锥或以 Ioffe 在一般 Banach 空间上的基本法锥的扩展来表示边际价格, 似乎不大可能获得如此极限分散类型的结果. 另一方面, 在 2.5.2 小节讨论的一些特殊情形中, 基本法向量可表示为其他一些具有变分结构的更原始的法向量的极限. 特别地, 在有限维空间和 Hilbert 空间上邻近法向量可以作到这一点, 这允许通过二次函数的特定扰动的最大化和最小化, 由此可提供极限边际价格的经济解释. 更多的细节和讨论请参看 Jofré [633]、Jofré 与 Rivera [635].

8.5.12 抽象版本

在第 8 章 (亦是全书) 的最后一节讨论 8.1~8.3 节中建立的第二福利定理扩展结果的进一步可能的对应版本和推广. 由于本书研究经济模型的方法主要基于变分分析的极点原理, 该节只简要探讨了可以方便应用特定版本的极点原理的一些情形.

首先考虑与 8.1~8.3 节中相同的模型并分析运用 2.5.3 节中建立的极点原理抽象版本的可能性. 读者可以发现, 尽管 8.1~8.3 节中所得到的某些结果 (特别与非线性价格和正性相关的) 需要 Fréchet- 型法向量和/或 Asplund 空间结构, 第二福利定理的大部分扩展在满足所给公理条件的具有恰当预法锥结构的其他 (有时是任意)Banach 空间情形同样成立. 本节也提到了 Jofré [633], Flåm 与 Jourani [453], 和 Flåm [452] 建立的由基于公理体系的法向量和次梯度来描述的第二福利定理的各种扩展. 与所论经济模型上第二福利定理的其他抽象扩展相比, 基于 Mordukhovich [920] 的 8.4.1 节的抽象结果似乎是最具一般性的.

8.5.13 进一步扩展

Samuelson 在他 1954 年的论文 [1189] 中, 在光滑的假设下首次提出有公共商品的福利经济模型. 他建立了 Hicks 和 Lange 的"基础"结果的公共商品版本. 这个基础结果指出, 公共商品的边际转换率等于个体边际替换率之和. 13 年以后, Foley[463] 对带有公共商品的凸经济学得出了 Arrow-Debreu 第二福利定理的相应

版本. 对凸经济学 Mas-Colell 适宜性框架下的公共商品福利理论的发展, 请参看近期的 De Simone、Graziano[326] 及其所录文献. 非凸模型的广义第二福利定理的各种结果可见 Khan 和 Vohra[673]、Khan [670, 671]、Flåm 和 Jourani [453]、Villar [1287, 1288], 以及其他研究者的论文.

由 8.4.2 节的讨论可知, 为没有公共商品的福利经济学建立的方法可以很容易地扩展到公共商品经济学中, 并保持 Samuelson 边际转换和替代率的基本结论; 参看 (8.52).

对其他模型 (如 Villar[1287] 中带有公共环境的模型, Makarov, Levin, Rubinov[829] 中带有直接配置的模型), Habte 的博士论文 [533] 中, 从通过极点原理处理的广义第二福利定理的角度, 有详细的讨论.

参考文献

1. Y. A. Abramovich and C. D. Aliprantis (2002), *An Invitation to Operator Theory*, American Mathematical Society, Providence, Rhode Island.
2. A. A. Agrachev and Y. L. Sachkov (2004), *Control Theory from the Geometric Viewpoint*, Springer, Berlin.
3. N. U. Ahmed (2005), Necessary conditions of optimality with discontinuous vector fields, *Nonlinear Funct. Anal. Appl.* **10**, 129–150.
4. N. U. Ahmed and K. L. Teo (1981), *Optimal Control of Distributed Parameter Systems*, Elsevier, New York.
5. N. U. Ahmed and X. Xiang (1997), Nonlinear uncertain systems and necessary conditions of optimality, *SIAM J. Control Optim.* **35**, 1755–1772.
6. N. U. Ahmed and X. Xiang (1997), Necessary conditions of optimality for differential inclusions in Banach spaces, *Nonlinear Anal.* **30**, 5437–5445.
7. V. M. Alekseev, V. M. Tikhomirov and S. V. Fomin (1987), *Optimal Control*, Consultants Bureau, New York.
8. A. D. Alexandrov (1939), Almost everywhere existence of the second differential of a convex function and some properties of convex surfaces connected with it, *Uchenye Zapiski Leningrad Gos. Univ., Ser. Math.* **6**, 3–35.
9. J. J. Alibert and J.-P. Raymond (1998), A Lagrange multiplier theorem for control problems with state constraints, *Numer. Funct. Anal. Optim.* **19**, 697–704.
10. C. D. Aliprantis, D. J. Brown and O. Burkinshaw (1990), *Existence and Optimality of Competitive Equilibria*, Springer, New York.
11. C. D. Aliprantis and O. Burkinshaw (1988), The fundamental theorems of welfare economics without proper preferences, *J. Math. Econ.* **17**, 41–54.
12. C. D. Aliprantis, M. Florenzano and R. Tourky (2005), Linear and non-linear price decentralization, *J. Econ. Theory*, **121**, 51–74
13. C. D. Aliprantis, P. K. Monteiro and R. Tourky (2004), Non-marketed options, non-existence of equilibria, and non-linear prices, *J. Econ. Theory* **114**, 345–357.
14. C. D. Aliprantis, R. Tourky and N. C. Yannelis (2001), A theory of value with non-linear prices: Equilibrium analysis beyond vector lattices, *J. Econ. Theory* **100**, 22–72.
15. K. Allali and T. Amahroq (1997), On openness and regularity of γ-paraconvex multifunctions, *Control Cybernet.* **26** (1997), 87–92.

16. T. Amahroq and N. Gadhi (2001), On the regularity conditions for vector programming problems, *J. Global Optim.* **21**, 435–443.
17. L. Ambrosio, O. Ascenzi and G. Buttazzo (1989), Lipschitz regularity of intergral functionals with highly discontinuous integrands, *J. Math. Anal. Appl.* **142**, 301–316.
18. R. M. Anderson (1988), The second welfare theorem with non-convex preferences, *Econometrica* **56**, 361–282.
19. T. S. Angell and A. Kirsch (1990), On the necessary conditions for optimal control of retarded systems, *Appl. Math. Optim.* **22**, 117–145.
20. M. Anitescu (2005), Global convergence of an elastic mode approach for a class of mathematical programs with complementarity constraints, *SIAM J. Control Optim.*, **16**, 120 – 145
21. N. Arada (2001), Minimax Dirichlet boundary control problem with state constraints, *Nonlinear Anal.* **46**, 653–673.
22. N. Arada, M. Bergounioux and J.-P. Raymond (2000), Minimax controls for uncertain distributed parabolic systems, *SIAM J. Control Optim.* **38**, 1481–1500.
23. N. Arada and J.-P. Raymond (1999), Optimality conditions for state-constrained Dirichlet boundary control problems, *J. Optim. Theory Appl.* **102**, 51–68.
24. N. Arada and J.-P. Raymond (2002), Dirichlet boundary control of semilinear parabolic equations, II: Problems with pointwise state constraints, *Appl. Math. Optim.* **45**, 145–167.
25. V. I. Arkin and V. L. Levin (1973), Convexity of the values of vector integrals, theorems on measurable selection, and variational problems, *Russian Math. Surveys* **27**, 21–85.
26. K. J. Arrow (1951), An extension of the basic theorems of classical welfare economics, in *Proceedings of the Second Berkeley Symposium on Mathematical Statistics and Probability*, pp. 507–532, University of California, Berkeley, California.
27. K. J. Arrow and L. Hurwicz (1960), Decentralization and computation in resource allocation, in *Essays in Economics and Econometrics in Honor of Harold Hotelling*, edited by R. W. Pfouts, pp. 34–104, University of North Carolina Press, Chapel Hills, North Carolina.
28. Z. Artstein (1994), First order approximations for differential inclusions, *Set-Valued Anal.* **2**, 7–17.
29. Z. Artstein (1995), A calculus for set-valued maps and set-valued evolution equations, *Set-Valued Anal.* **3**, 213–261.
30. Z. Artstein and V. Gaitsgory (2000), The value function of singularly perturbed control systems, *Appl. Math. Optim.* **41**, 425–445.
31. Z. Artstein and C. C. Popa (2004), Orlicz-Young structures of Young measures, *Acta Appl. Math.* **80**, 1–33.
32. A. V. Arutyunov (2000), *Optimality Conditions: Abnormal and Degenerate Problems*, Kluwer, Dordrecht, The Netherlands.
33. A. V. Arutyunov and S. M. Aseev (1997), Investigation of the degeneracy phenomenon of the maximum principle for optimal control problems with state constraints, *SIAM J. Control Optim.* **35**, 930–952.
34. A. V. Arutyunov, S. M. Aseev and V. I. Blagodatskikh (1994), Necessary conditions of the first order in the problem of optimal control of a differential inclusion with phase constraints, *Math. Sbornik* **79**, 117–139.

35. A. V. Arutyunov and A. F. Izmailov (2005), Sensitivity analysis for cone-constrained optimization problems under relaxed constraint qualifications, *Math. Oper. Res.* **30**, 333-353.
36. A. V. Arutyunov and A. F. Izmailov (2006), Directional stability theorem and directional metric regularity, *Math. Oper. Res.*, **31**, 526–543.
37. A. V. Arutyunov and F. L. Pereira (2006), Second-order necessary conditions of optimality for problems without a priori normality assumptions, *Math. Oper. Res.*, **31**, 1–12.
38. A. V. Arutyunov and R. B. Vinter (2004), A simple 'finite approximations' proof of the Pontryagin maximum principle under reduced differentiability hypotheses, *Set-Valued Anal.* **12**, 5–24.
39. S. M. Aseev (1991), Smooth approximations of differential inclusions and the time-optimality problem, *Proc. Steklov Inst. Math.* **200**, 27–34.
40. S. M. Aseev (1997), A method of smooth approximations in the theory of necessary optimality conditions for differential inclusions, *Izvestia: Mathematics* **61**, 235–258.
41. S. M. Aseev (2001), Extremal problems for differential inclusions with phase constraints, *Proc. Steklov Inst. Math.* **233**, 1–63.
42. L. T. Ashchepkov and O. V. Vasiliev (1975), On optimality of singular controls in Goursat-Darboux systems, *USSR Comput. Maths. Math. Phys.* **15**, 1157–1167.
43. E. Asplund (1968), Fréchet differentiability of convex functions, *Acta Math.* **121**, 31–47.
44. H. Attouch (1984), *Variational Convergence for Functions and Operators*, Pitman, Boston, Massachusetts.
45. H. Attouch and H. Ruani (1993), Stability results for Ekeland's ε-variational principle and cone extremal solutions, *Math. Oper. Res.* **18**, 173–201.
46. H. Attouch and R. J-B. Wets (1991), Quantitative stability of variational systems, I: The epigraphical distance, *Trans. Amer. Math. Soc.* **338**, 695–729.
47. H. Attouch and R. J-B. Wets (1993), Quantitative stability of variational systems, III: ε-approximate solutions, *Math. Progr.* **61**, 197–214.
48. J.-P. Aubin (1981), Contingent derivatives of set-valued maps and existence of solutions to nonlinear inclusions and differential inclusions, in *Mathematical Analysis and Applications*, edited by L. Nachbin, pp. 159–229, Academic Press, New York.
49. J.-P. Aubin (1984), Lipschitz behavior of solutions to convex minimization problems, *Math. Oper. Res.* **9**, 87–111.
50. J.-P. Aubin and A. Cellina (1984), *Differential Inclusions*, Springer, Berlin.
51. J.-P. Aubin and I. Ekeland (1976), Estimates of the duality gap in nonconvex programming, *Math. Oper. Res.* **1**, 225–245.
52. J.-P. Aubin and I. Ekeland (1984), *Applied Nonlinear Analysis*, Wiley, New York.
53. J.-P. Aubin and H. Frankowska (1987), On inverse function theorems for set-valued maps, *J. Math. Pures Appl.* **66**, 71–89.
54. J.-P. Aubin and H. Frankowska (1990), *Set-Valued Analysis*, Birkhäuser, Boston, Massachusetts.
55. R. J. Aumann (1965), Integrals of set-valued functions, *J. Math. Anal. Appl.* **12**, 1–12.

56. A. AUSLENDER (1978), Differential stability in nonconvex and nondifferentiable programming, *Math. Progr. Study* **10**, 29–41.
57. A. AUSLENDER (1984), Stability in mathematical programming with nondifferentiable data, *SIAM J. Control Optim.* **22**, 239–254.
58. A. AUSLENDER AND R. COMINETTI (1991), A comparable study of multifunction differentiability with applications in mathematical programming, *Math. Oper. Res.* **16**, 240–258.
59. A. AUSLENDER AND J.-P. CROUZEIX (1988), Global regularity theorems, *Math. Oper. Res.* **13**, 243–253.
60. A. AUSLENDER AND M. TEBOULLE (2003), *Asymptotic Cones and Functions in Optimization and Variational Inequalities*, Springer, New York.
61. D. AUSSEL, J.-N. CORVELLEC AND M. LASSONDE (1995), Mean value property and subdifferential criteria for lower semicontinuous functions, *Trans. Amer. Math. Soc.* **347**, 4147–4161.
62. D. AUSSEL, J.-N. CORVELLEC AND M. LASSONDE (1995), Nonsmooth constrained optimization and multidimensional mean value inequalities, *SIAM J. Optim.* **9**, 690–706.
63. D. AUSSEL, A. DANIILIDIS AND L. THIBAULT (2004), Subsmooth sets: Functional characterizations and related concepts, *Trans. Amer. Math. Soc.* **357**, 1275–1302.
64. E. P. AVAKOV (1988), Necessary conditions of minima for nonregular problems in Banach spaces: Maximum principle for abnormal problems of optimal control, *Proc. Steklov Inst. Math.* **185**, 3–29.
65. E. R. AVAKOV, A. A. AGRACHEV AND A. V. ARUTYUNOV (1992), The level set of a smooth mapping in a neighborhood of a singular point, *Math. Sbornik* **73**, 455–466.
66. J. AVELIN (1997), *Differential Calculus cfor Multifunctions and Nonsmooth Functions*, Ph.D. dissertation, Department of Mathematics, Uppsala University, Sweden.
67. J. AVELIN (2000), Differential calculus for complex-valued multifunctions, *J. Appl. Anal.* **6**, 47–76.
68. V. I. AVERBUKH AND O. G. SMOLYANOV (1968), The various definitions of the derivative in linear topological spaces, *Russian Math. Surveys* **23**, 67–113.
69. D. AZÉ AND J.-N. CORVELLEC (2004), Characterizations of error bounds for lower semicontinuous functions on metric spaces, *Control Optim. Calc. Var.*, **10**, 409-425.
70. D. AZÉ, J.-N. CORVELLEC AND R. E. LUCCHETTI (2002), Variational pairs and applications to stability in nonsmooth analysis, *Nonlinear Anal.* **49**, 643–670.
71. D. AZÉ AND J.-B. HIRIART-URRUTY (2002), Optimal Hoffman type estimates in eigenvalue and semidefinite inequality constraints, *J. Global Optim.* **24**, 133–147.
72. R. L. BAIRE (1905), *Lecons sur les Fonctions Discontinues*, Gauthier-Villars, Paris.
73. A. V. BALAKRISHNAN (1965), Optimal control problems in Banach spaces, *SIAM J. Control* **3**, 152–180.
74. A. V. BALAKRISHNAN (1981), *Applied Functional Analysis*, 2nd edition, Springer, New York.
75. E. J. BALDER (2000), Lectures on Young measure theory and its applications to economics, *Rens. Istit. Mat. Univ. Trieste* **31**, 1–69.

76. S. BANACH (1932), *Théorie des Opérations Linéaires*, Monografje Matematyczne, Warsaw.
77. B. BANK, J. GUDDAT, D. KLATTE, B. KUMMER AND K. TAMMER (1982), *Non-Linear Parametric Optimization*, Birkhaüser, Basel.
78. H. T. BANKS (1968), Necessary conditions for control problems with variable time lags, *SIAM J. Control* **6**, 9–47.
79. H. T. BANKS AND G. A. KENT (1972), Control of functional differential equations of retarded and neutral type to target sets in function spaces, *SIAM J. Control* **10**, 567–591.
80. H. T. BANKS AND K. KUNISCH (1989), *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, Massachusetts.
81. H. T. BANKS AND A. MANITIUS (1974), Applications of abstract variational theory to hereditary systems – a survey, *IEEE Trans. Autom. Control* **19**, 524–533.
82. V. BARBU (1993), *Analysis and Control of Nonlinear Infinite Dimensional Systems*, Academic Press, Boston, Massachusetts.
83. V. BARBU, I. LASIECKA AND R. TRIGGIANI (2005), Boundary stabilization of Navier-Stokes equations, *Mem. Amer. Math. Soc.*, to appear.
84. V. BARBU AND T. PRECUPANU (1986), *Convexity and Optimization in Banach Spaces*, D. Reidel, Dordrecht, The Netherlands.
85. M. BARDI AND I. CAPUZZO DOLCETTA (1997), *Optimal Control and Viscosity Solutions of Hamilton-Jacobi Equations*, Birkhäuser, Boston, Massachusetts.
86. N. E. BARRON AND R. JENSEN (1991), Optimal control and semicontinuous viscosity solutions, *Proc. Amer. Math. Soc.* **113**, 397–402.
87. T. BAŞAR AND P. BERNHARD (1995), *H_∞-Optimal Control and Related Minimax Design Problems*, 2nd edition, Birkhäuser, Boston, Massachusetts.
88. H. BAUSCHKE, J. M. BORWEIN AND W. LI (1999), Strong conical hull intersection property, bounded linear regularity, Jameson's property (G), and error bounds in convex optimization, *Math. Progr.* **86**, 135–160.
89. M. S. BAZARAA, J. J. GOODE AND M. Z. NASHED (1974), An the cone of tangents with applications to mathematical programming, *J. Optim. Theory Appl.* **13**, 389–426.
90. C. R. BECTOR, S. CHANDRA AND J. DUTTA (2004), *Principles of Optimization Theory*, Alpha Science International Publishers, UK.
91. E. M. BEDNARCZUK (2003), Order-Lipschitzian properties of multifunctions with applications to stability of efficient points, *Control Cybernet.* **32**, 491–502.
92. G. BEER (1993), *Topologies on Closed and Convex Sets*, Kluwer, Dordrecht, The Netherlands.
93. S. BELLAASSALI AND A. JOURANI (2004), Necessary optimality conditions in multiobjective dynamic optimization, *SIAM J. Control Optim.* **42**, 2043–2061.
94. R. BELLMAN AND K. L. COOKE (1963), *Differential-Difference Equations*, Academic Press, New York.
95. R. BELLMAN, I. GLICKSBERG AND O. GROSS (1956), On the "bang-bang" control problem, *Quart. Appl. Math.* **14**, 11–18.
96. H. BENABDELLAH (2000), Existence of solutions to the nonconvex sweeping process, *J. Diff. Eq.* **164**, 286-295.
97. H. BENABDELLAH, C. CASTAING, A. SALVADORI AND A SYAM (1996), Nonconvex sweeping process, *J. Appl. Anal.* **2**, 217–240.

98. M. BENAIM, J. HOFBAUER AND S. SORIN (2005), Stochastic approximations and differential inclusions, *SIAM J. Control Optim.* **44**, 328–348.
99. J. BENOIST (1994), Approximation and regularization of arbitrary sets in finite dimensions, *Set-Valued Anal.* **2**, 95–115.
100. A. BENSOUSSAN (1988), *Perturbation Methods in Optimal Control*, Wiley, New York.
101. A. BENSOUSSAN, G. DA PRATO, M. DELFOUR, S. K. MITTER (1992–93), *Representation and control of infinite-dimensional systems*, published in two volumes, Birkhäuser, Boston, Massachusetts.
102. A. BEN-TAL AND J. ZOWE (1985), Directional derivatives in nonsmooth optimization, *J. Optim. Theory App.* **47**, 483–490.
103. M. BERGOUNIOUX AND D. TIBA (1996), General optimality conditions for constrained convex control problems, *SIAM J. Control Optim.* **34**, 698–711.
104. M. BERGOUNIOUX AND F. TRÖLTZSCH (1996), Optimality conditions and generalized bang-bang principle for a state-constrained semilinear parabolic problem, *Numer. Func. Anal. Optim.* **17**, 517–536.
105. L. D. BERKOVITZ (1974), *Optimal Control Theory*, Springer, New York.
106. L. D. BERKOVITZ (1976), A penalty function proof of the maximum principle, *Appl. Math. Optim.* **2**, 291–303.
107. H. BERLIOCCHI AND J. M. LASRY (1973), Principe de Pontryagin pour des systémes régis par une equation différentielle multivoque, *C. R. Acad. Sci. Paris* **277**, 1103–1105.
108. F. BERNARD AND L. THIBAULT (2004), Prox-regularity of functions and sets in Banach spaces, *Set-Valued Anal.* **12**, 25–47.
109. F. BERNARD AND L. THIBAULT (2005), Prox-regular functions in Hilbert spaces, *J. Math. Anal. Appl.* **303**, 1–14.
110. F. BERNARD AND L. THIBAULT (2005), Uniform prox-regularity of functions and epigraphs in Hilbert spaces, *Nonlinear Anal.* **60**, 187–207.
111. D. P. BERTSEKAS (1999), *Nonlinear Programming*, Athena Scientific, Boston, Massachusetts.
112. D. P. BERTSEKAS AND A. E. OZDAGLAR (2002), Pseudonormality and a Lagrange multiplier theory for constraint optimization, *J. Optim. Theory Appl.* **114**, 287–343.
113. D. BESSIS, Y. S. LEDYAEV AND R. B. VINTER (2001) Dualization of the Euler and Hamiltonian inclusions, *Nonlinear Anal.* **43**, 861–882.
114. R.-M. BIANCHINI AND M. KAWSKI (2003), Needle variations that cannot be summed, *SIAM J. Control Optim.* **42**, 218–238.
115. J. BIRGE AND L. QI (1995), Subdifferential convergence in stochastic programming, *SIAM J. Optim.* **5**, 436–453.
116. E. BISHOP AND R. R. PHELPS (1963), The support functionals of convex sets, in *Convexity*, edited by V. Klee, vol. VII of *Proceedings of Symposia in Pure Mathematics*, pp. 27–35, American Mathematical Society, Providence, Rhode Island.
117. V. I. BLAGODATSKIKH (1984), The maximum principle for differential inclusions, *Proc. Steklov Inst. Math.* **166**, 23–43.
118. V. I. BLAGODATSKIKH AND A. F. FILIPPOV (1986), Differential inclusions and optimal control, *Proc. Steklov Inst. Math.* **169**, 199–259.
119. G. A. BLISS (1946), *Lectures on the Calculus of Variations*, University of Chicago Press, Chicago, Illinois.

120. N. A. Bobylev, S. V. Emel'yanov and S. K. Korovin (2004), Convexity of images of convex sets under smooth maps, *Comput. Math. Model.* **15**, 213–222.
121. N. N. Bogolyubov (1930), Sur quelques methodes nouvelles dans le calculus des variations, *Ann. Math. Pura Appl.* **7**, 249–271.
122. J. Bolte, A. Daniilidis and A. Lewis (2006), The Lojasiewicz inequality for nonsmooth subanalytic functions with applications to subgradient dynamical systems, *SIAM J. Optim.* **17**, 1205–1223.
123. J. Bolte, A. Daniilidis and A. Lewis (2006), A nonsmooth Morse-Sard theorem for subanalytic functions, *J. Math. Anal. Appl.*, **321**, 729–740.
124. V. G. Boltyanskii (1958), The maximum principle in the theory of optimal processes, *Dokl. Akad. Nauk SSSR* **119**, 1070–1073.
125. V. G. Boltyanskii (1968), Local section method in the theory of optimal processes, *Diff. Eq.* **4**, 2166–2183.
126. V. G. Boltyanskii (1973), The maximum principle for problems of optimal steering, *Diff. Eq.* **9**, 1363–1370.
127. V. G. Boltyanskii (1973), *Optimal Control of Discrete Systems*, Nauka, Moscow.
128. V. G. Boltyanskii (1994), The maximum principle – how it came to be?, Report No. 526, Math. Institut, Tech. Univ. München, München, Germany.
129. V. G. Boltyanskii, R. V. Gamkrelidze and L. S. Pontryagin (1956), On the theory of optimal processes, *Dokl. Akad. Nauk SSSR* **110**, 7–10.
130. O. Bolza (1913), Über den anormalen fall beim Lagrangeschen und Mayerschen problem mit gemischten behingugen und variabeln endpunkten, *Mathematische Annalen* **74**, 430–446.
131. J. F. Bonnans and E. Casas (1992), A boundary Pontryagin's principle for the optimal control of state-constrained elliptic systems, *Int. Ser. Numer. Math.* **107**, 241–249.
132. J. F. Bonnans, R. Cominetti and A. Shapiro (1999), Second-order optimality conditions based on parabolic second-order tangent sets, *SIAM J. Optim.* **9**, 466–492.
133. J. F. Bonnans and A. Shapiro (2000), *Perturbation Analysis of Optimization Problems*, Springer, New York.
134. J. F. Bonnans and A. Sulem (1995), Pseudopower expansion of solutions of generalized equations and constraint optimization problems, *Math. Progr.* **70**, 123–148.
135. J.-M. Bonnisseau and B. Cornet (1988), Valuation equilibrium and Pareto optimum in non-convex economies, *J. Math. Econ.* **17**, 293–308.
136. D. Borwein, J. M. Borwein and X. Wang (1996), Approximate subgradients and coderivatives in $I\!R^n$, *Set-Valued Anal.* **4**, 375–398.
137. J. M. Borwein (1986), Stability and regular points of inequality systems, *J. Optim. Theory Appl.* **48**, 9–52.
138. J. M. Borwein (1987), Epi-Lipschitz-like sets in Banach spaces: Theorems and examples, *Nonlinear Anal.* **11**, 1207–1217.
139. J. M. Borwein, M. Fabian, I. Kortezov and P. D. Loewen (2001), The range of the gradient of a continuously differentiable bump, *J. Nonlinear Convex Anal.* **2**, 1–19.
140. J. M. Borwein, M. Fabian and P. D. Loewen (2002), The shape of the range of a C^1 smooth bump in infinite dimensions, *Israel J. Math.* **132**, 239–251.

141. J. M. BORWEIN AND S. P. FITZPATRICK (1995), Weak* sequential compactness and bornological limit derivatives, *J. Convex Anal.* **2**, 59–68.
142. J. M. BORWEIN AND S. P. FITZPATRICK (1995), Characterization of Clarke subgradients among one-dimensional multifunctions, in *Proc. Optimization Miniconference II*, edited by B. M. Glover and V. Jeyakumar, pp. 61–64, University of New South Wales, Sydney, Australia.
143. J. M. BORWEIN AND S. P. FITZPATRICK (2001), Duality inequalities and sandwiched functions, *Nonlinear Anal.* **46**, 365–380.
144. J. M. BORWEIN, S. P. FITZPATRICK AND R. GIRGENSOHN (2003), Subdifferentials whose graphs are not norm×weak* closed, *Canad. Math. Bull.* **46**, 538–545.
145. J. M. BORWEIN, S. P. FITZPATRICK AND J. R. GILES (1987), The differentiability of real functions on normed linear spaces using generalized subgradients, *J. Math. Anal. Appl.* **128**, 512–534.
146. J. M. BORWEIN AND J. R. GILES (1987), The proximal normal formula in Banach spaces, *Trans. Amer. Math. Soc.* **302**, 371–381.
147. J. M. BORWEIN AND A. D. IOFFE (1996), Proximal analysis in smooth spaces, *Set-Valued Anal.* **4**, 1–24.
148. J. M. BORWEIN AND A. JOFRÉ (1998), A nonconvex separation property in Banach spaces, *Math. Meth. Oper. Res.* **48**, 169–179.
149. J. M. BORWEIN AND A. S. LEWIS (2000), *Convex Analysis and Nonlinear Optimization: Theory and Examples*, Springer, New York.
150. J. M. BORWEIN, Y. LUCET AND B. S. MORDUKHOVICH (2000), Compactly epi-Lipschitzian convex sets and functions in normed spaces, *J. Convex Anal.* **7**, 375–393.
151. J. M. BORWEIN, B. S. MORDUKHOVICH AND Y. SHAO (1999), On the equivalence of some basic principles of variational analysis, *J. Math. Anal. Appl.* **229**, 228–257.
152. J. M. BORWEIN, W. B. MOORS AND X. WANG (2001), Generalized subdifferentials: A Baire categorical approach, *Trans. Amer. Math. Soc.* **353**, 3875–3893.
153. J. M. BORWEIN AND D. NOLL (1994), Second order differentiability of convex functions in Banach spaces, *Trans. Amer. Math. Soc.* **342**, 43–81.
154. J. M. BORWEIN AND D. PREISS (1987), A smooth variational principle with applications to subdifferentiability and differentiability of convex functions, *Trans. Amer. Math. Soc.* **303**, 517–527.
155. J. M. BORWEIN AND H. M. STRÓJWAS (1985), Tangential approximations, *Nonlinear Anal.* **9**, 1347–1366.
156. J. M. BORWEIN AND H. M. STRÓJWAS (1986), Proximal analysis and boundaries of closed sets in Banach spaces, I: Theory, *Canad. J. Math.* **38**, 431–452.
157. J. M. BORWEIN AND H. M. STRÓJWAS (1987), Proximal analysis and boundaries of closed sets in Banach spaces, II: Applications, *Canad. J. Math.* **39**, 428–472.
158. J. M. BORWEIN, J. S. TREIMAN AND Q. J. ZHU (1998), Necessary conditions for constrained optimization problems with semicontinuous and continuous data, *Trans. Amer. Math. Soc.* **350**, 2409–2429.
159. J. M. BORWEIN, J. S. TREIMAN AND Q. J. ZHU (1999), Partially smooth variational principles and applications, *Nonlinear Anal.* **35**, 1031–1059.

160. J. M. BORWEIN AND Q. J. ZHU (1996), Viscosity solutions and viscosity subderivatives in smooth Banach spaces with applications to metric regularity, *SIAM J. Control Optim.* **34**, 1568–1591.

161. J. M. BORWEIN AND Q. J. ZHU (1997), Variational analysis in nonreflexive spaces and applications to control problems with L^1 perturbations, *Nonlinear Anal.* **28**, 889–915.

162. J. M. BORWEIN AND Q. J. ZHU (1998), Limiting convex examples for nonconvex subdifferential calculus, *J. Convex Anal.* **5**, 221–235.

163. J. M. BORWEIN AND Q. J. ZHU (1999), A survey of subdifferential calculus with applications, *Nonlinear Anal.* **38**, 687–773.

164. J. M. BORWEIN AND Q. J. ZHU (2005) *Techniques of Variational Analysis*, Springer, New York.

165. J. M. BORWEIN AND D. M. ZHUANG (1988), Verifiable necessary and sufficient conditions for openness and regularity of set-valued and single-valued maps, *J. Math. Anal. Appl.* **134**, 441–459.

166. P. BOSCH, A. JOURANI AND R. HENRION (2004), Sufficient conditions for error bounds and applications, *Appl. Math. Optim.* **50**, 161–181.

167. G. BOULIGAND (1930), Sur les surfaces dépourvues de points hyperlimits, *Ann. Soc. Polon Math.* **9**, 32–41.

168. G. BOULIGAND (1932), *Introduction à la Géométrie Infinitésimale Directe*, Gauthier-Villars, Paris.

169. R. D. BOURGIN (1983), *Geometric Aspects of Convex Sets with the Radon-Nicodým Property*, Springer, New York, 1993.

170. M. BOUNKHEL (2004), Scalarization of the normal Fréchet regularity of set-valued mappings, *New Zealand J. Math.* **33**, 1–18.

171. M. BOUNKHEL AND A. JOFRÉ (2004), Subdifferential stability of the distance function and its applications to nonconvex economics and equilibrium, *J. Nonlinear Convex Anal.* **5**, 331–347.

172. M. BOUNKHEL AND L. THIBAULT (2002), On various notions of regularity of sets in nonsmooth analysis, *Nonlinear Anal.* **48**, 223–246.

173. M. BOUNKHEL AND L. THIBAULT (2005), Further characterizations of regular sets in Hilbert spaces and their applications to nonconvex sweeping processes, *J. Nonlinear Convex Anal.*, **6**, 359–374.

174. K. E. BRENNAN, S. L. CAMPBELL AND L. R. PRETZOLD (1989), *Numerical Solution of Initial Value Problems in Differential-Algebraic Equations*, North-Holland, New York.

175. A. BRESSAN (1990), Upper and lower semicontinuous differential inclusions: A unified approach, in *Nonlinear Controllability and Optimal Control*, edited by H. J. Sussmann, pp. 21-31, Marcel Dekker, New York.

176. O. A. BREZHNEVA AND A. A. TRET'YAKOV (2003), Optimality conditions for degenerate extremum problems with equality constraints, *SIAM J. Control Optim.* **42**, 729–745.

177. H. BRÉZIZ AND W. STRAUSS (1973), Semilinear second-order elliptic equations in L^1, *J. Math. Soc. Japan* **25**, 565–590.

178. J. BRINKHUIS AND V. M. TIKHOMIROV (2005), *Optimization: Insights and Applications*, Princeton University Press, Princeton, New Jersey.

179. A. BRØNDSTED AND R. T. ROCKAFELLAR (1965), On the subdifferentiability of convex functions, *Proc. Amer. Math. Soc.* **16**, 605–611.

180. L. E. J. BROUWER (1910), Über eineindeutige, stetige transformationen von flächen in sich, *Math. Ann.* **69**, 176–180.

181. D. J. BROWN (1991), Equilibrium analysis with nonconvex technologies, in *Handbook of Mathematical Economics*, Vol. 4, edited by W. Hildenbrand and H. Sonnenschein, North-Holland, Amsterdam, The Netherlands.
182. A. BRUCKNER (1994), *Differentiation of Real Functions*, American Mathematical Society, Providence, Rhode Island.
183. F. BUCCI (1992), A Dirichlet boundary control problem for the strongly damped wave equation, *SIAM J. Control Optim.* **30**, 1092–1100.
184. B. M. BUDAK, E. M. BERKOVICH AND E. N. SOLOVIEVA (1969), Difference approximations in optimal control problems, *SIAM J. Control* **7**, 18–31.
185. B. M. BUDAK AND F. P. VASILIEV (1975), *Some Computational Aspects of Optimal Control Problems*, Moscow University Press, Moscow.
186. R. S. BURACHIK AND A. N. IUSEM (2006), *Set-Valued Mappings and Enlargements of Monotone Operators*, SIAM Publications, Philadelphia, Pennsylvania.
187. R. S. BURACHIK, A. N. IUSEM AND B. F. SVAITER (1997), Enlargement of monotone operators with applications to variational inequalities, *Set-Valued Anal.* **5**, 159–180.
188. J. V. BURKE (1991), Calmness and exact penalization, *SIAM J. Control Optim.* **29**, 493–497.
189. J. V. BURKE (1991), An exact penalization viewpoint of constrained optimization, *SIAM J. Control Optim.* **29**, 968–998.
190. J. V. BURKE AND S. DENG (2002), Weak sharp minima revisted, I: Basic theory, *Control Cybernet.* **31**, 439–469.
191. J. V. BURKE AND S. DENG (2005), Weak sharp minima revisted, II: Application to linear regularity and error bounds, *Math. Progr.*, **104**, 235–261.
192. J. V. BURKE AND M. C. FERRIS (1993),On the Clarke subdifferential of the distance function of a closed set, Weak sharp minima in mathematical programming, *SIAM J. Control Optim.* **31**, 1340–1359.
193. J. V. BURKE, M. C. FERRIS AND M. QUIAN (1992), *J. Math. Anal. Appl.* **166**, 199–213.
194. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2000), Optimizing matrix stability, *Proc. Amer. Math. Soc.* **129**, 1635–1642.
195. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2001), Optimal stability and eigenvalue multiplicity, *Found. Comput. Math.* **1**, 205–225.
196. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2002), Approximating subdifferentials by random sampling of gradients, *Math. Oper. Res.* **27**, 567–584.
197. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2003), Robust stability and a criss-cross algorithm for pseudospectra, *IMA J. Numerical Anal.* **23**, 1–17.
198. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2004), Variational analysis of the abscissa mapping for polynomials via the Gauss-Lucas theorem, *J. Global Optim.* **28**, 259–268.
199. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2004), A robust gradient sampling algorithm for nonsmooth, nonconvex optimization, *SIAM J. Optim.* **15**, 751–779.
200. J. V. BURKE, A. S. LEWIS AND M. L. OVERTON (2005), Variational analysis of functions of the roots of polynomials, *Math. Progr.*,**104**, 263–292.
201. J. V. BURKE AND D. R. LUKE (2003), Variational analysis applied to the problem of optical phase retrieval, *SIAM J. Control Optim.* **42**, 576–595.

202. J. V. Burke and M. L. Overton (1992), in *Nonsmooth Optimization Methods and Applications*, edited by F. Giannessi, pp. 19–29, Gordon and Breach, Philadelphia, Pennsylvania.
203. J. V. Burke and M. L. Overton (2000), Variational analysis of the abscissa mapping for polynomials, *SIAM J. Control Optim.* **39**, 1651–1676.
204. J. V. Burke and M. L. Overton (2001), Variational analysis of non-Lipschitz spectral functions, *Math. Progr.* **90**, 317–351.
205. J. V. Burke and P. Tseng (1996), A unified analysis of Hoffman's bound via Fenchel duality, *SIAM J. Optim.* **6**, 265–282.
206. J. A. Burns (2003), Nonlinear distributed parameter control systems with non-normal linearizations: Applications and approximations, *Front. Appl. Math.* **27**, 17–53.
207. M. Bustos (1994), Epsilon gradients pour functions localements Lipschitziennes et applications, *Numer. Funct. Anal. Optim.* **15**, 435–453.
208. A. G. Butkovsky (1963), Necessary and sufficient optimality conditions for impulse control systems, *Autom. Remote Control* **24**, 1056–1064.
209. A. G. Butkovsky (1969), *Distributed Control Systems*, Elsevier, New York.
210. A. G. Butkovsky, A. I. Egorov and K. A. Lurie (1968), Optimal control of distributed systems (a survey of Soviet publications), *SIAM J. Control* **6**, 437–476.
211. A. G. Butkovsky and A. J. Lerner (1960), Optimal control systems with distributed parameters, *Autom. Remote Control* **21**, 472–477.
212. G. Buttazzo and G. Dal Maso (1982), Γ-convergence and optimal control problems, *J. Optim. Theory Appl.* **38**, 385–407.
213. G. Buttazzo, M. Giaquinta and S. Hildebrandt (1998), *One-Dimensional Variational Problems: An Introduction*, Oxford University Press, New York.
214. F. Camilli and M. Falcone (1996), Approximation of optimal control problems with state constraints: Estimates and applications, in *Nonsmooth Analysis and Geometric Methos in Deterministic Optimal Control*, edited by B. S. Mordukhovich and H. J. Sussmann, pp. 23–58, Springer, New York.
215. P.-M. Cannarsa and G. Da Prato (1991), Second-order Hamilton-Jacobi equations in infinite dimensions, *SIAM J. Control Optim.* **29**, 474–492.
216. P.-M. Cannarsa and H. Frankowska (1991), Some characterizations of optimal trajectories in control theory, *SIAM J. Control Optim.* **29**, 1322–1347.
217. P.-M. Cannarsa and C. Sinestrari (2004), *Semiconvex Functions, Hamilton-Jacobi Equations, and Optimal Control*, Birkhäuser, Boston, Massuchusetts.
218. M. Cannon, C. Cullum and E. Polak (1970), *Theory of Optimal Control and Mathematical Programming*, McGraw-Hill, New York.
219. M. J. Cánovas, A. L. Dontchev, M. A. López and J. Parra (2005), Metric regularity of semi-infinite constraint systems, *Math. Progr.*, to appear.
220. I. Capuzzo Dolcetta (1983), On a discrete approximation of the Hamilton-Jacobi bequation of dynamic programming, *Appl. Math. Optim.* **10**, 367–377.
221. I. Capuzzo Dolcetta and M. Falcone (1989), Discrete dynamic programming and visocity solutions of the Bellman equation, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **6**, 161–184.
222. C. Carathéodory (1935), *Variationsrechung und Partielle Differentialgleichungen erster Ordnung*, B. G. Teubner, Leipzig and Berlin, 1935.

223. O. Cârjă (1984), The time optimal problem for boundary-distributed control systems, *Bull. Unione Mat. Italiana* **3-B**, 563–581.
224. D. A. Carlson, A. Haurie and A. Leizarowitz (1991), *Infinite Horizon Optimal Control*, Springer, Berlin.
225. E. Casas (1997), Pontryagin's principle for state-constrained boundary control problems of semilinear parabolic equations, *SIAM J. Control Optim.* **35**, 1297–1327.
226. E. Casas, J.-P. Raymond and H. Zidani (2000), Pontryagin's principle for local solutions of control problems with mixed control-state constraints, *SIAM J. Control Optim.* **39**, 1182–1203.
227. E. Casas and J. Yong (1995), Maximum principle for state-constrained optimal control problems governed by quasilinear elliptic equations, *Diff. Integ. Eq.* **8**, 1–18.
228. C. Castaing, A. Salvadori and L. Thibault (2001), Functional evolution equations governed by nonconvex sweeping process, *J. Nonlinear Convex Anal.* **2**, 217–241.
229. C. Castaing and M. Valadier (1977), *Convex Analysis and Measurable Multifunctions*, Springer, New York.
230. E. Cavazzuti, M. Pappalardo and M. Passacantando (2002), Nash equilibria, variational inequalities, and dynamical systems, *J. Optim. Theory Appl.* **114**, 491–505.
231. A. Cellina (1988), On the set of solutions to Lipschitzian differential inclusions, *Diff. Integ. Eq.* **1**, 459–500.
232. A. Cellina (2003), The classical problem of the calculus of variations in the autonomous case: Relaxation and Lipschitzianity of solutions, *Trans. Amer. Math. Soc.* **356**, 415–426.
233. A. Cernea (2001), Some second-order necessary conditions for non-convex hyperbolic differential inclusions problems, *J. Math. Anal. Appl.* **253**, 616–639.
234. A. Cernea (2005), Second-order necessary conditions for differential-difference inclusion problems, *Nonlinear Anal.* **62**, 963–974.
235. L. Cesari (1983), *Optimization – Theory and Applications*, Springer, New York.
236. R. W. Chaney (1987), Second-order directional derivatives for nonsmooth functions, *J. Math. Anal. Appl.* **128**, 495–511.
237. A. V. Cherkaev (2000), *Variational Methods for Structural Optimization*, Springer, New York.
238. F. L. Chernousko (1994), *State Estimation for Dynamic Systems*, CRC Press, Boca Raton, Florida.
239. A. A. Chikrii (1997), *Conflict-Controlled Processes*, Kluwer, Dordrecht, The Netherlands.
240. I. Chryssochoos and R. B. Vinter (2003), Optimal control on manifolds: A dynamic programming approach, *J. Math. Anal. App.* **287**, 118–140.
241. E. N. Chukwu (1992), *Stability and Time-Optimal Control of Hereditary Systems*, Academic Press, Boston, Massachusetts.
242. M. Ciligot-Travain and S. Traore (2002), On subgradients of spectral functions, *J. Convex Anal.* **9**, 401–414.
243. F. H. Clarke (1973), *Necessary Conditions for Nonsmooth Problems in Optimal Control and the Calculus of Variations*, Ph.D. dissertation, Department of Mathematics, University of Washington, Seattle.

244. F. H. CLARKE (1975), Generalized gradients and applications, *Trans. Amer. Math. Soc.* **205**, 247–262.
245. F. H. CLARKE (1975), The Euler-Lagrange differential inclusion, *J. Diff. Eq.* **19**, 80–90.
246. F. H. CLARKE (1975), Admissible relaxation in variational and control problems, *J. Math. Anal. Appl.* **51**, 557–576.
247. F. H. CLARKE (1976), Optimal solutions to differential inclusions, *J. Optim. Theory Appl.* **19**, 469–478.
248. F. H. CLARKE (1976), The generalized problem of Bolza, *SIAM J. Control Optim.* **14**, 682–699.
249. F. H. CLARKE (1976), A new approach to Lagrange multipliers, *Math. Oper. Res.* **2**, 165–174.
250. F. H. CLARKE (1976), The maximum principle under minimal hypotheses, *SIAM J. Control Optim.* **14**, 1078–1091.
251. F. H. CLARKE (1976), Necessary conditions for a general control problem, in *Calculus of Variations and Control Theory*, edited by D. L. Russel, pp. 257–278, Academic Press, New York.
252. F. H. CLARKE (1976), On the inverse function theorem, *Pacific J. Math.* **64**, 97–102.
253. F. H. CLARKE (1977), Extremal arcs and extended Hamiltonian systems, *Trans. Amer. Math. Soc.* **231**, 349–367.
254. F. H. CLARKE (1980), The Erdmann condition and Hamiltonian inclusions in optimal control and the calculus of variations, *Can. J. Math.* **32**, 494–509.
255. F. H. CLARKE (1983), *Optimization and Nonsmooth Analysis*, Wiley, New York.
256. F. H. CLARKE (1987), Hamiltonian analysis of the generalized problem of Bolza, *Trans. Amer. Math. Soc.* **301**, 385–400.
257. F. H. CLARKE (1989), *Methods of Dynamic and Nonsmooth Optimization*, SIAM Publications, Philadelphia, Pennsylvania.
258. F. H. CLARKE (1993), A decoupling principle in the calculus of variations, *J. Math. Anal. Appl.* **172**, 92–105.
259. F. H. CLARKE (1993), An indirect method in the calculus of variations, *Trans. Amer. Math. Soc.* **336**, 655-673.
260. F. H. CLARKE (2003), Necessary conditions in optimal control: A new approach, *Math. Progr.* **97**, 71–89.
261. F. H. CLARKE (2005), Necessary conditions in dynamic optimization, *Mem. Amer. Math. Soc.* **173**, No. 816.
262. F. H. CLARKE AND Y. S. LEDYAEV (1994), Mean value inequality in Hilbert spaces, *Trans. Amer. Math. Soc.* **344**, 307–324.
263. F. H. CLARKE, Y. S. LEDYAEV, E. D. SONTAG AND A. I. SUBBOTIN (1997), Asymptotic controllability implies feedback stabilization, *IEEE Trans. Autom. Control* **42**, 1394–1407.
264. F. H. CLARKE, Y. S. LEDYAEV, R. J. STERN AND P. R. WOLENSKI (1995), Qualitative properties of trajectories of control systems: A survey, *J. Dynam. Control Systems* **1**, 1–47.
265. F. H. CLARKE, Y. S. LEDYAEV, R. J. STERN AND P. R. WOLENSKI (1998), *Nonsmooth Analysis and Control Theory*, Springer, New York.
266. F. H. CLARKE, Y. S. LEDYAEV AND P. R. WOLENSKI (1995), Proximal analysis and minimization principles, *J. Math. Anal. Appl.* **196**, 722–735.

267. F. H. CLARKE AND P. D. LOEWEN (1986), The value function in optimal control: Sensitivity, controllability and time-optimality, *SIAM J. Control Optim.* **24**, 243–263.
268. F. H. CLARKE AND P. D. LOEWEN (1987), State constraints in optimal control: A case study in proximal normal analysis, *SIAM J. Control Optim.* **25**, 1440–1456.
269. F. H. CLARKE AND R. J. STERN (2003), State constrained feedback stabilization, *SIAM J. Control Optim.* **42**, 422–441.
270. F. H. CLARKE, R. J. STERN AND P. R. WOLENSKI (1993), Subgradient criteria for monotonicity and the Lipschitz condition, *Canad. J. Math.* **45**, 1167–1183.
271. F. H. CLARKE, R. J. STERN AND P. R. WOLENSKI (1995), Proximal smoothness and the lower-C^2 property, *J. Convex Anal.* **2**, 117–144.
272. F. H. CLARKE AND R. B. VINTER (1989), Applications of optimal multiprocesses, *SIAM J. Control Optim.* **25**, 1048–1071.
273. F. H. CLARKE AND R. B. VINTER (1989), Optimal multiprocesses, *SIAM J. Control Optim.* **25**, 1072–1091.
274. F. H. CLARKE AND G. G. WATKINS (1986), Necessary conditions, controllability and the value function for differential-difference inclusions, *Nonlinear Anal.* **10**, 1155–1179.
275. F. H. CLARKE AND P. R. WOLENSKI (1991), The sensitivity of optimal control problems to time delay, *SIAM J. Control Optim.* **29**, 1176–1215.
276. F. H. CLARKE AND P. R. WOLENSKI (1996), Necessary conditions for functional differential inclusions, *Appl. Math. Optim.* **34**, 34–51.
277. G. COLOMBO AND V. V. GONCHAROV (1999), The sweeping processes without convexity, *Set-Valued Anal.* **7**, 357–374.
278. G. COLOMBO AND V. V. GONCHAROV (2001), Variational inequalities and regularity properties of closed sets in Hilbert spaces, *J. Convex Anal.* **8**, 197–221.
279. G. COLOMBO AND A. MARIGONDA (2006), Differentiability properties for a class of nonconvex functions, *Calc. Var. Partial Diff. Eq.*, **25**, 1–31.
280. G. COLOMBO AND P. R. WOLENSKI (2004), Variational analysis for a class of minimal time functions in Hilbert spaces, *J. Convex Anal.* **11**, 335–361.
281. F. COLONIUS (1982), The maximum principle for relaxed hereditary differential systems with function space end conditions, *SIAM J. Control Optim.* **20**, 695–712.
282. R. COMINETTI (1990), Metric regularity, tangent cones, and second-order optimality conditions, *Appl. Math. Optim.* **21**, 265–287.
283. R. COMINETTI AND R. CORREA (1990), A generalized second order derivative in nonsmooth optimization, *SIAM J. Optim.* **28**, 789–809.
284. R. CONTI (1968), Time-optimal solution of a linear evolution equation in Banach spaces, *J. Optim. Theory Appl.* **2**, 277–284.
285. B. CORNET (1981), Regular properties of tangent and normal cones, CEREMADE Publication 8130, Universiteé de Paris IX "Dauphine".
286. B. CORNET (1986), The second welfare theorem in nonconvex economies, CORE discussion paper No. 8630.
287. B. CORNET (1988), General equilibrium theorem and increasing returns: Presentation, *J. Math. Econ.* **17**, 103–118.

288. B. Cornet (1990), Marginal cost pricing and Pareto optimality, in *Essays in Honor of Edmond Malinvaud*, edited by P. Champsaur, Vol. 1, pp. 14–53, MIT Press, Cambridge, Massachusetts.
289. B. Cornet and M.-O. Czarnecki (1999), Smooth normal approximations of epi-Lipschitz subset in $\mathbb{R}^n$, *SIAM J. Control Optim.* **37**, 710–730.
290. B. Cornet and M.-O. Czarnecki (2001), Existence of generalized equilibria: Necessary and sufficient conditions, *Nonlinear Anal.* **44**, 555–574.
291. R. Correa, P. Gajardo and L. Thibault (2005), Subdifferential representation formula and subdifferential criteria for the behavior of nonsmooth functions, *Nonlinear Anal.*, to appear.
292. R. Correa, A. Jofré and L. Thibault (1994), Subdifferential monotonicity as characterization of convex functions, *Numer. Funct. Anal. Optim.* **15**, 1167–1183.
293. R. W. Cottle, F. Giannessi and J.-L. Lions, ed. (1980), *Variational Inequalities and Complementarity Problems*, Wiley, New York.
294. R. W. Cottle, J.-S. Pang and R. E. Stone (1992), *The Linear Complementarity Problem*, Academic Press, Boston, Massachusetts.
295. M. G. Crandall, L. C. Evans and P.-L. Lions (1984), Some properties of viscosity solutions, *Trans. Amer. Math. Soc.* **282**, 487–502.
296. M. G. Crandall, H. Ishii and P.-L. Lions (1992), User's guide to viscosity solutions of second-order partial differential equations, *Bull. Amer. Math. Soc.* **27**, 1–67.
297. M. G. Crandall and P.-L. Lions (1983), Viscosity solutions of Hamilton-Jacobi equations, *Trans. Amer. Math. Soc.* **277**, 1–42.
298. B. D. Craven (1994), Convergence of discrete approximations for constrained minimization, *J. Austral. Math. Soc.* **35**, 1–12.
299. B. D. Craven (1995), *Control and Optimization*, Chapman and Hall, London, UK.
300. B. D. Craven and D. V. Luu (1997), A method for establishing optimality conditions for a nonconvex vector-valued minimax problem, *J. Optim. Theory Appl.* **95**, 295–304.
301. G. P. Crespi, D. La Torre and M. Rocca (2003), Second-order mollified derivatives and optimization, *J. Nonlinear Convex Anal.* **3**, 437-454.
302. J. Cullum (1969), Discrete approximations to continuous optimal control problems, *SIAM J. Control* **7**, 32–49.
303. J. Cullum (1972), Finite-dimensional approximations of state constrained continuous optimal control problems, *SIAM J. Control* **10**, 649–670.
304. M.-O. Czarnecki and L. Rifford (2006), Approximation and regularization of Lipschitz functions: Convergence of the gradients, *Trans. Amer. Math. Soc.*, **358**, 4467–4520.
305. L. Dai (1989), *Singular Control Systems*, Springer, Berlin.
306. P. Daniele and A. Maugeri (2001), On dynamical equilibrium problems and variational inequalities, in *Equilibrium Problems: Nonsmooth Optimization and Variational Inequality Models*, edited by F. Giannessi, A. Maugeri and P. Pardalos, pp. 59–69, Kluwer, Dordrecht, The Netherlands.
307. J. M. Danskin (1967), *The Theory of Min-Max and Its Application to Weapon Allocations Problems*, Springer, New York.
308. F. S. De Blasi, G. Pianigiani and A. A. Tolstonogov (2004), A Bogolyubov type theorem with a nonconvex constraint in Banach spaces, *SIAM J. Control Optim.* **43**, 466-476.

309. G. Debreu (1951), The coefficient of resource utilization, *Econometrica* **19**, 273–292.
310. G. Debreu (1959), *Theory of Values*, Yale University Press, New Haven, Connecticut.
311. G. Debreu (1970), Economies with a finite set of equilibria, *Econometrica* **38**, 387–392.
312. E. De Giorgi, A. Marino and M. Tosques (1980), Problemi di evoluzione in spazi metric e curve di massima pendenza, *Atti. Accad. Naz. Lincei Rend. Cl. Sci. Fis. Mat. Natur.* **68**, 180–187.
313. M. Degiovanni, A. Marino and M. Tosques (1985), Evolution equations with lack of convexity, *Nonlinear Anal.* **9**, 1401–1443.
314. K. Deimling (1992), *Multivalued Differential Equations*, De Gruyter, Berlin.
315. J. Demmel (1987), The condition number and the distance to the nearest ill-posed problem, *Numerische Math.* **51**, 251–289.
316. S. Dempe (2002), *Foundations of Bilevel Programming*, Kluwer, Dordrecht, The Netherlands.
317. S. Dempe (2003), Annotated bibliography on bilevel programming and mathematical programs with equilibrium constraints, *Optimization* **52**, 333–359.
318. V. F. Demyanov (2005), *Extremality Conditions and Calculus of Variations*, Vysshaya Shkola, Moscow.
319. V. F. Demyanov and V. N. Malozemov (1974), *Introduction to Minimax*, Wiley, New York.
320. V. F. Demyanov amd A. M. Rubinov (1980), On quasidifferentiable functionals, *Dokl. Akad. Nauk SSSR* **250**, 21–25.
321. V. F. Demyanov and A. M. Rubinov (1995), *Constructive Nonsmooth Analysis*, Peter Lang, Frankfurt, Germany.
322. V. F. Demyanov and A. M. Rubinov, eds. (1995), *Quasidifferentiability and Related Topics* (2000), Kluwer, Dordrecht, The Netherlands.
323. Z. Denkowski, S. Migórski and N. S. Papageorgiou (2004), *An Introduction to Nonlinear Analysis and Its Applications*, published in two volumes, Kluwer, Boston, Massachusetts.
324. D. Dentcheva and W. Römisch (2000), Differential stability of two-stage stochastic programs, *SIAM J. Optim.* **11**, 87–112.
325. D. Dentcheva and A. Ruszczyński (2004), Optimization with stochastic dominance constraints, *SIAM J. Optim.* **14**, 548-566.
326. A. De Simone and M. G. Graziano (2004), The pure theory of public goods: The case of many commodities, *J. Math. Econ.* **40**, 847–868.
327. E. N. Devdariani and Y. S. Ledyaev (1999), Maximum principle for implicit control systems, *Appl. Math. Optim.* **41**, 79–103.
328. R. Deville (1994), Stability of subdifferentials of nonconvex functions in Banach spaces, *Set-Valued Anal.* **2**, 141–157.
329. R. Deville and E. M. El Haddad (1996), The viscosity subdifferential of the sum of two functions in Banach spaces, I: First order case, *J. Convex Anal.* **3**, 295–308.
330. R. Deville, G. Godefroy and V. Zizler (1993), A smooth variational principle with applications to Hamilton-Jacobi equations in infinite dimensions, *J. Funct. Anal.* **111**, 197–212.
331. R. Deville, G. Godefroy and V. Zizler (1993), *Smoothness and Renorming in Banach Spaces*, Wiley, New York.

332. J. DIESTEL (1975), *Geometry of Banach Spaces–Selected Topics*, Springer, Berlin.
333. J. DIESTEL (1984) *Sequences and Series in Banach Spaces*, Springer, New York.
334. J. DIESTEL AND J. J. UHL, JR. (1977), *Vector Measures*, American Mathematical Society, Providence, Rhode Island.
335. U. DINI (1878), *Fondamenti per la Teoria delle Funzioni di Variabili Reali*, Pisa, Italy.
336. A. V. DMITRUK (2002), An nonlocal Lyusternik estimate and its applications to control systems with sliding modes, in *Nonlinear Control Systems 2001*, edited by A. B. Kurzhanskii and A. L. Fradkov, Vol. 2, pp. 1061–1064, Elsevier, Exeter, UK.
337. A. V. DMITRUK, A. A. MILYUTIN AND N. P. OSMOLOVSKII (1980), Lyusternik's theorem and the theory of extrema, *Russian Math. Surveys* **35**, 11–51.
338. B. DOITCHINOV AND V. VELIOV (1993), Parameterizations of integrals of set-valued mappings and applications, *J. Math. Anal. Appl.* **179**, 483–499.
339. S. DOLECKI (1980), A general theory of necessary optimality conditions, *J. Math. Anal. Appl.* **78**, 267–308.
340. S. DOLECKI (1982), Tangency and differentiation: Some applications of convergence theory, *Ann. Mat. Pura Appl.* **130**, 223–255.
341. S. DOLECKI AND S. ROLEWICZ (1979), Exact penalties for local minima, *SIAM J. Control Optim.* **17**, 596–606.
342. T. DONCHEV (1997), Lower semicontinuous differential inclusions: One-sided Lipschitz approach, *Coll. Math.* **74**, 177–184.
343. T. DONCHEV (2005), Approximation of solution sets for optimal control problems, in *Large-Scale Scientific Computations*, edited by I. Lirkov, S. Margenov and J. Wasniewski, Lecture Notes Comp. Sci., Springer, Berlin.
344. T. DONCHEV AND E. FARKHI (1998), Stability and Euler approximation of one-sided Lipschitz differential inclusions, *SIAM J. Control Optim.* **36**, 780–796.
345. T. DONCHEV AND E. FARKHI (1999), Euler approximation of discontinuous one-sided Lipschitz convex differential inclusions, in *Calculus of Variations and Differential Equations*, edited by A. Ioffe, S. Reich and I. Shafrir, pp. 101–118, Chapman and Hall/CRC Press, Boca Raton, Florida.
346. T. DONCHEV, E. FARKHI AND B. S. MORDUKHOVICH (2007), Discrete approximations, relaxation, and optimization of one-sided Lipschitzian differential inclusions in Hilbert spaces, *J. Diff. Eqs.* **243**, 301–328.
347. A. L. DONTCHEV (1981), Error estimates for s discrete approximation to constrained control problems, *SIAM J. Numer. Anal.* **18**, 500–514.
348. A. L. DONTCHEV (1983), *Perturbations, Approximations and Sensitivity Analysis of Optimal Control Systems*, Springer, Berlin.
349. A. L. DONTCHEV (1988), Equivalent perturbations and approximations in optimal control, in *International Series of Numerical Mathematics* **84**, pp. 43–54, Birkhäuser, Basel, Switzerland.
350. A. L. DONTCHEV (1995), Characterization of Lipschitz stability in optimization, in *Recent Developments in Well-Posed Variational Problems and Related Topics*, edited by R. Lucchetti and J. Revalski, pp. 95–116, Kluwer, Dordrecht, The Netherlands.

351. A. L. DONTCHEV (1995), Implicit function theorems for generalized equations, *Math. Progr.* **70**, 91–106.
352. A. L. DONTCHEV (1996), The Graves theorem revisited, *J. Convex Anal.* **3**, 45–54.
353. A. L. DONTCHEV (1996), Discrete approximations in optimal control, in *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control*, edited by B. S. Mordukhovich and H. J. Sussmann, pp. 59–81, Springer, New York.
354. A. D. DONTCHEV AND E. M. FARKHI (1989), Error estimates for discretized differential inclusions, *Computing* **41**, 349–358.
355. A. L. DONTCHEV AND W. W. HAGER (1993), Lipschitzian stability in nonlinear control and optimization, *SIAM J. Control Optim.* **31**, 569–603.
356. A. L. DONTCHEV AND W. W. HAGER (1994), Implicit functions, Lipschitz maps and stability in optimization, *Math. Oper. Res.* **19**, 753–768.
357. A. L. DONTCHEV AND W. W. HAGER (1998), A new approach to Lipschitz continuity in state constrained optimal control, *Syst. Control Lett.* **35**, 137–143.
358. A. L. DONTCHEV AND W. W. HAGER (2001), The Euler approximation in state constrained optimal control, *Math. Comp.* **70**, 173–203.
359. A. L. DONTCHEV AND F. LEMPIO (1992), Difference methods for differential inclusions: A survey, *SIAM Rev.* **34**, 263–294.
360. A. L. DONTCHEV AND A. S. LEWIS (2005), Perturbations and metric regularity, *Set-Valued Anal.*,**13**, 417–438.
361. A. L. DONTCHEV, A. S. LEWIS AND R. T. ROCKAFELLAR (2003), The radius of metric regularity, *Trans. Amer. Math. Soc.* **355**, 493–517.
362. A. L. DONTCHEV AND B. S. MORDUKHOVICH (1983), Relaxation and well-posedness of nonlinear optimal processes, *Syst. Control Lett.* **3**, 177–179.
363. A. L. DONTCHEV, M. QUINCAMPOIX AND N. ZLATEVA (2006), Aubin criterion for metric regularity, *J. Convex Anal.*, **13**, 281-297.
364. A. L. DONTCHEV AND R. T. ROCKAFELLAR (1996), Characterizations of strong regularity for variational inequalities over polyhedral convex sets, *SIAM J. Optim.* **7**, 1087–1105.
365. A. L. DONTCHEV AND R. T. ROCKAFELLAR (2001), Ample parameterization of variational inclusions, *SIAM J. Optim.* **12**, 170–187.
366. A. L. DONTCHEV AND R. T. ROCKAFELLAR (2004), Regularity and conditioning of solution mappings in variational analysis, *Set-Valued Anal.* **12**, 79–109.
367. A. L. DONTCHEV AND T. ZOLEZZI (1993), *Well-Posed Optimization Problems*, Springer, New York.
368. A. DOUGLIS (1961), The continuous dependence of generalized solutions of non-linear partial differential equations upon initial data, *Comm. Pure Appl. Math.* **14**, 267–284.
369. A. Y. DUBOVITSKII AND A. A. MILYUTIN (1963), Extremum problems in the presence of constraints, *Dokl. Akad. Nauk SSSR* **149**, 759–762.
370. A. Y. DUBOVITSKII AND A. A. MILYUTIN (1965), Extremum problems in the presence of restrictions, *USSR Comput. Maths. Math. Phys.* **5**, 1–80.
371. N. DUNFORD AND J. T. SCHWARTZ (1958), *Linear Operators, I: General Theory*, Interscience, New York.
372. J. C. DUNN (1995), Second order optimality conditions in sets of L^∞ functions with range in a polyhedron, *SIAM J. Control Optim.* **33**, 1603–1635.

373. M. DÜR, R. HORST AND M. LOCATELLI (1998), Necessary and sufficient conditions for convex maximization revisited, *J. Math. Anal. Appl.* **217**, 637–649.
374. J. DUTTA (2005), Necessary optimality conditions and saddle points for approximate optimization in Banach spaces, *TOP: Spanish J. Oper. Res.*, **13**, 127–143.
375. J. DUTTA (2005), Optimality conditions for maximizing a locally Lipschitz function, *Optimization* **54**, 377–389.
376. J. DUTTA (2005), Generalized derivatives and nonsmooth optimization: A finite-dimensional tour, *TOP: Spanish J. Oper. Res.*, **13**.127–143.
377. J. DUTTA AND S. DEMPE (2005), Bilevel programming with convex lower level problems, in *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, edited by S. Dempe and V. V. Kalashnikov, Springer, Berlin.
378. J. DUTTA AND C. TAMMER (2006), Lagrangian conditions for vector optimization on Banach spaces, *Math. Meth. Oper. Res.*,**64**, 521–540.
379. A. N. DYUKALOV (1983), *Problems of Applied Mathematical Economics*, Nauka, Moscow.
380. Z. DZALILOV, A. F. IVANOV AND A. M. RUBINOV (2001), Difference inclusions with delay of economic growth, *Dynam. Syst. Appl.* **10**, 283–293.
381. A. EBERHARD (2001), Prox-regularity and subjects, in *Optimization and Related Topics*, edited by A. Rubinov and B. Glover, *Applied Optimization* **47**, pp. 237–313, Kluwer, Dordrecht, The Netherlands.
382. A. EBERHARD AND M. NYBLOM (1998), Jets, generalized convexity, proximal normality and differences of functions, *Nonlinear Anal.* **34**, 319–360.
383. A. EBERHARD, M. NYBLOM AND D. RALPH (1998), Applying generalized convexity notions to jets, in *Generalized Convexity, Generalized Monotonicity: Recent Results*, edited by J.-P. Crouzeix, J. E. Martinez-Legaz and M. Volle, pp. 111–157, Kluwer, Dordrecht, The Netherlands.
384. A. EBERHARD AND C. E. M. PEARCE (2005), A comparison of two approaches to second-order subdifferentiability concepts with application to optimality conditions, in *Optimization and Control with Applications*, edited by L. Qi, K. L. Teo and X. Yang, pp. 36–100, Springer, Berlin.
385. A. EBERHARD, C. E. M. PEARCE AND D. RALPH (1998), A comparative study of jets, graphical derivatives and coderivatives, RMIT Research Report, No. 8, Melbourne, Australia.
386. A. EBERHARD, R. SIVAKUMARAN AND R. WENCZEL (2005), On the variational behavior of the subhessians of the Lasry-Lions envelope, *J. Convex Anal.*, to appear.
387. A. EBERHARD AND R. WENCZEL (2009), Some sufficient optimality conditions in nonsmooth analysis, *SIAM J. Optim.* **20**, 251–296.
388. C. ECKART AND G. YOUNG (1936), The approximation of one matrix by another of lower rank, *Psychometrika* **1**, 211–218.
389. J. E. EDMOND AND L. THIBAULT (2002), Inclusions and integration of subdifferentials, *J. Nonlinear Convex Anal.* **3**, 411–434.
390. J. E. EDMOND AND L. THIBAULT (2005), Relaxation of an optimal control problem involving a perturbed sweeping process, *Math. Progr.*,**104**, 347–373.
391. A. I. EGOROV (1967), Necessary optimality conditions for distributed parameter systems, *SIAM J. Control* **5**, 352–408.
392. A. I. EGOROV (1978), *Optimal Control of Heat and Diffusion Processes*, Nauka, Moscow.

393. Y. V. EGOROV (1963), Certain problems in optimal control theory, *USSR Comput. Maths. Math. Phys.* **3**, 1209–1232.
394. Y. V. EGOROV (1964), Some necessary conditions for optimality in Banach spaces, *Math. Sbornik* **64**, 79–101.
395. K. J. EISENHART (2003), *Multiobjective Optimal Control Problems with Endpoint and State Constraints*, Ph.D. dissertation, Departments of Mathematics, Western Michigan University, Kalamazoo, Michigan.
396. I. EKELAND (1972), Sur les problémes variationnels, *C. R. Acad. Sci. Paris* **275**, 1057–1059.
397. I. EKELAND (1974), On the variational principle, *J. Math. Anal. Appl.* **47**, 324–353.
398. I. EKELAND (1974), Une estimation a priori en programmation non convexe, *C. R. Acad. Sci. Paris* **277**, 149–151.
399. I. EKELAND (1979), Nonconvex minimization problems, *Bull. Amer. Math. Soc.* **1**, 432–467.
400. I. EKELAND AND G. LEBOURG (1976), Generic Fréchet differentiability and perturbed optimization problems in Banach spaces, *Trans. Amer. Math. Soc.* **224**, 193–216.
401. I. EKELAND AND R. TEMAM (1976), *Convex Analysis and Variational Problems*, North-Holland, Amsterdam, The Netherlands.
402. B. EL ABDOUNI AND L. THIBAULT (1992), Lagrange multipliers for Pareto nonsmooth programming problems in Banach spaces, *Optimization* **26**, 277–285.
403. L. E. ÉLSGOLTS (1955), *Qualitative Methods in Mathematical Analysis* (1955), GITTL, Moscow.
404. L. E. ÉLSGOLTS AND S. B. NORKIN (1973), *Introduction to the Theory and Applications of Differential Equations with Deviating Arguments*, Academic Press, New York.
405. M. EPELMAN AND R. M. FREUND (2002), A new condition measure, preconditioners, and relations between different measures of conditioning for conic linear systems, *SIAM J. Optim.* **13**, 627–655.
406. I. I. EREMIN (1966), The penalty method in convex programming, *Soviet Math. Dokl.* **8**, 458–462.
407. Y. M. ERMOLIEV, V. P. GULENKO AND T. I. TZARENKO (1978), *Finite Difference Method in Optimal Control Problems*, Naukova Dumka, Kiev.
408. Y. M. ERMOLIEV, V. I. NORKIN AND R. J-B. WETS (1995), The minimization of semicontinuous functions: Mollifier subgradients, *SIAM J. Control Optim.* **33**, 149–167.
409. E. ERNST AND M. THÉRA (2005), A converse to the Eidelgeit theorem in real Hilbert spaces, *Bull. Sci. Math.* **129**, 381–397.
410. E. ERNST AND M. THÉRA (2006), Global maximum of a convex function: Necessary and sufficient conditions, *J. Convex Anal.* **13**, 687-694.
411. L. EULER (1744), *Methodus Inveniendi Curvas Lineas Maximi Minimive Proprietate Gaudentes Sive Solution Problematis Viso Isoperimetrici Latissimo Sensu Accepti*, Lausanne; reprinted in *Opera Omnia*, Ser. 1, Vol. 24, 1952.
412. U. G. EVTUSHENKO (1982), *Methods of Solving Extremal Problems and Their Application in Optimization Systems*, Nauka, Moscow.
413. M. FABIAN (1986), Subdifferentials, local ε-supports and Asplund spaces, *J. London Math. Soc.* **34**, 568–576.

414. M. FABIAN (1988), On classes of subdifferentiability spaces of Ioffe, *Nonlinear Anal.* **12**, 568–576.
415. M. FABIAN (1989), Subdifferentiability and trustworthiness in the light of a new variational principle of Borwein and Preiss, *Acta Univ. Carolina, Ser. Math. Phys.* **30**, 51–56.
416. M. FABIAN (1997), *Gâteaux Differentiability of Convex Functions and Topology. Weak Asplund Spaces*, Wiley, New York.
417. M. FABIAN, P. HÁJEK AND J. VANDERWERFF (1996), On smooth variational principles in Banach spaces, *J. Math. Anal. Appl.* **197**, 153–172.
418. M. FABIAN, P. D. LOEWEN AND B. S. MORDUKHOVICH (2006), Subdifferential calculus in Asplund generated spaces, *J. Math. Anal. Appl.*,**322**, 787–795.
419. M. FABIAN AND B. S. MORDUKHOVICH (1998), Smooth variational principles and characterizations of Asplund spaces, *Set-Valued Anal.* **6**, 381–406.
420. M. FABIAN AND B. S. MORDUKHOVICH (1999), Separable reduction and supporting properties of Fréchet-like normals in Banach spaces, *Canad. J. Math.* **51**, 26–48.
421. M. FABIAN AND B. S. MORDUKHOVICH (2002), Separable reduction and extremal principles in variational analysis, *Nonlinear Anal.* **49**, 265–292.
422. M. FABIAN AND B. S. MORDUKHOVICH (2003), Sequential normal compactness versus topological normal compactness in variational analysis, *Nonlinear Anal.* **54**, 1057–1067.
423. M. FABIAN AND N. V. ZHIVKOV (1985). A characterization of Asplund spaces with help of local ε-supports of Ekeland and Lebourg, *C. R. Acad. Bulgare Sci.* **38**, 671–674.
424. F. FACCHINEI AND J.-S. PANG (2003), *Finite-Dimensional Variational Inequalities and Complementarity Problems*, published in two volumes, Springer, New York.
425. M. FALCONE (1994), Discrete time high-order schemes for viscosity solutions of the Hamilton-Jacobi equation, *Numer. Math.* **67**, 315–344.
426. L.-T. FAN AND C.-S. WANG (1964), *The Discrete Maximum Principle: A Study of Multistage Systems Optimization*, Wiley, New York.
427. H. O. FATTORINI (1964), Time-optimal control of solutions of operational differential equations, *SIAM J. Control* **2**, 54–59.
428. H. O. FATTORINI (1968), Boundary control problems, *SIAM J. Control* **6**, 349–385.
429. H. O. FATTORINI (1987), A unified theory of necessary conditions for nonlinear nonconvex control systems, *Appl. Math. Optim.* **15**, 141–185.
430. H. O. FATTORINI (1996), Optimal control with state constraints for semilinear distributed parameter systems, *J. Optim. Theory Appl.* **88**, 25–59.
431. H. O. FATTORINI (1997), Nonlinear infinite dimensional optimal control problems with state constraints and unbounded control sets, *Rend. Instit. Mat. Univ. Trieste* **28**, 127–146.
432. H. O. FATTORINI (1999), *Infinite Dimensional Optimization and Control Theory*, Cambridge University Press, Cambridge, UK.
433. H. O. FATTORINI (2001), The maximum principle for control systems described by linear parabolic equations, *J. Math. Anal. Appl.* **259**, 630–651.
434. H. O. FATTORINI AND H. FRANKOWSKA (1991), Necessary conditions for infinite dimensional control problems, *Math. Contr. Signal Syst.* **4**, 41–67.

435. H. O. Fattorini and T. Murphy (1994), Optimal problems for nonlinear parabolic boundary control problems: Dirichlet boundary conditions, *Diff. Integ. Eq.* **7**, 1367–1388.
436. H. O. Fattorini and T. Murphy (1994), Optimal problems for nonlinear parabolic boundary control systems, *SIAM J. Control Optim.* **32**, 1577-1596.
437. H. Federer (1959), Curvature measures, *Trans. Amer. Math. Soc.* **93**, 418–491.
438. R. P. Fedorenko (1970), Maximum principle for differential inclusions, *USSR Comput. Maths. Math. Phys.* **10**, 57–68.
439. R. P. Fedorenko (1971), Maximum principle for differential inclusions (necessity), *USSR Comput. Maths. Math. Phys.* **11**, 885–893.
440. A. A. Feldbaum (1953), Optimal process in systems of automatic control, *Autom. Telemekh.***14**, No. 5.
441. H. Fenchel (1951), *Convex Cones, Sets and Functions*, Lecture Notes, Princeton University, Princeton, New Jersey.
442. P. de Fermat (1638), Letters to M. Mersenne and J. Roberval, Toulouse, France.
443. M. M. A. Ferreira, F. A. C. C. Fontes and R. B. Vinter (1999), Nondegenerate necessary conditions for nonconvex optimal control problems with state constraints, *J. Math. Anal. Appl.* **233**, 116–129.
444. M. M. A. Ferreira and R. B. Vinter (1994), When is the maximum principle for state-constrained problems degenerate?, *J. Math. Anal. Appl.* **187**, 432–467.
445. M. C. Ferris and O. L. Mangasarian (1993), Error bounds and strong upper semicontinuity for monotone affine variational inequalities, *Ann. Oper. Res.* **47**, 293–305.
446. M. C. Ferris and J.-S. Pang (1997), Engineering and economic applications of complementarity problems, *SIAM Rev.* **39**, 669–713.
447. A. V. Fiacco (1983), *Introduction to Sensitivity and Stability Analysis in Nonlinear Programming*, Academic Press, New York.
448. A. V. Fiacco and G. P. McCormick (1968), *Nonlinear Programming: Sequential Unconstrained Minimization Techniques*, Wiley, New York.
449. A. F. Filippov (1962), On certain questions in the theory of optimal control, *SIAM J. Control* **1**, 76–84.
450. A. F. Filippov (1988), *Differential Equations with Discontinuous Right-Hand Sides*, Kluwer, Dordrecht, The Netherlands.
451. S. P. Fitzpatrick (1980), Metric projections and the differentiability of distance functions, *Bull. Austral. Math. Soc.* **22**, 291–312.
452. S. D. Flåm (2006), Upward slopes and inf-convolutions, *Math. Oper. Res.*, **31**, 188–198.
453. S. D. Flåm and A. Jourani (2006), Prices and Pareto optima, *Optimization* **55**, 611-625.
454. M. L. Flegel (2005), *Constraint Qualifications and Stationarity Concepts for Mathematical Programs with Equilibrium Constraints*, Ph.D. Dissertation, Facult. Math., Univ. Würzburg, Germany.
455. M. L. Flegel and C. Kanzow (2003), A Fritz John approach to first order optimality conditions for mathematical programs with equilibrium constraints, *Optimization* **52**, 277–296.

456. M. L. FLEGEL AND C. KANZOW (2005), On M-stationarity points for mathematical programs with equilibrium constraints, *J. Math. Anal. Appl.* **310**, 286–302.
457. M. L. FLEGEL, C. KANZOW AND J. V. OUTRATA (2007), Optimality conditions for disjunctive programs with equilibrium constraints, *Set-Valued Anal.* **15**, 139–162.
458. W. H. FLEMING AND H. M. SONER (1993), *Controlled Markov Processes and Viscosity Solutions*, Springer, New York.
459. M. FLORENZANO (2003), *General Equilibrium Analysis: Existence and Optimality Properties of Equilibria*, Kluwer, Dordrecht, The Netherlands.
460. M. FLORENZANO, P. GOURDEL AND A. JOFRÉ (2006), Supporting weakly Pareto optimal allocations in infinite dimensional nonconvex economies, *J. Econ. Theory* **29**, 549–564.
461. F. FLORES-BAZÁN (1997), On minima of the difference functions, *J. Optim. Theory Appl.* **93**, 525–531.
462. F. FLORES-BAZÁN AND W. OETTLI (2001), Simplified optimality conditions for minimizing the difference of vector-valued convex functions, *J. Optim. Theory Appl.* **108**, 571–586.
463. D. FOLEY (1967), Resource allocation and the public sector, *Yale Econ. Essays* **7**, 43–98.
464. H. FRANKOWSKA (1985), Necessary conditions for the Bolza problem, *Math. Oper. Res.* **10**, 361–366.
465. H. FRANKOWSKA (1987), The maximum principle for an optimal control to a differential inclusion with endpoint constraints, *SIAM J. Control Optim.* **25**, 145–157.
466. H. FRANKOWSKA (1987), An open mapping principle for set-valued maps, *J. Math. Anal. Appl.* **127**, 172–180.
467. H. FRANKOWSKA (1989), Higher order inverse function theorems, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **6**, 283–304.
468. H. FRANKOWSKA (1989), Contingent cones to reachable sets of control systems, *SIAM J. Control Optim.* **27**, 170–198.
469. H. FRANKOWSKA (1990), Some inverse mappings theorems, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **7**, 183–234.
470. H. FRANKOWSKA (1990), A priori estimates for operational differential inclusions, *J. Diff. Eq.* **84**, 100–128.
471. H. FRANKOWSKA (1993), Lower semicontinuous solutions of Hamilton-Jacobi-Bellman equations, *SIAM J. Control Optim.* **31**, 257–272.
472. H. FRANKOWSKA (2005), Optimal synthesis via superdifferentials of value functions, *Control Cybernet.* **34** (2005), No. 3.
473. M. FRÉCHET (1911), Sur la notion de différentielle, *C. R. Acad. Sci. Paris* **152**, 845–847.
474. R. A. FREEMAN AND P. V. KOKOTOVIĆ (1996), *Robust Nonlinear Control Design: State-Space and Lyapunov Techniques*, Birkhäuser, Boston, Massachusetts.
475. R. M. FREUND AND J. R. VERA (1999), Some characterizations and properties of the 'distance to ill-posedness' and the condition measure of a conic linear system, *Math. Progr.* **86**, 225–260.
476. A. FRIEDMAN (1964), Optimal control for hereditary processes, *J. Math. Anal. Appl.* **15**, 396–414.
477. A. FRIEDMAN (1967), Optimal control in Banach spaces, *J. Math. Anal. Appl.* **19**, 35–55.

478. A. Friedman (1982), *Variational Principles and Free-Boundary Problems*, Wiley, New York.
479. A. Friedman (1987), Optimal control for parabolic variational inequalities, *SIAM J. Control Optim.* **25**, 482–497.
480. M. Fukushima and J.-S. Pang (2005), Quasi-variational inequalities, generalized Nash equilibria, and multi-leader-follower games, *Comput. Management Sci.* **1**, 21–56.
481. A. V. Fursikov (2000), *Optimal Control of Distributed Systems. Theory and Applications*, American Mathematical Society, Prividence, Rhode Island.
482. P. Fusek, D. Klatte and B. Kummer (2002), Examples and couterexamples in Lipschitz analysis, *Control Cybernet.* **31**, 471–492.
483. R. Gabasov and S. V. Churakova (1968), Necessary optimality conditions in time-lag systems, *Autom. Remote Control* **29**, 37–54.
484. R. Gabasov and F. M. Kirillova (1966), On the extension of the maximum principle by L. S. Pontryagin to discrete systems, *Autom. Remote Control* **27**, 1878–1882.
485. R. Gabasov and F. M. Kirillova (1974), *Maximum Principle in the Theory of Optimal Processes*, Nauka i Tekhnika, Minsk, Belarus.
486. R. Gabasov and F. M. Kirillova (1976), *Qualitative Theory of Optimal Processes*, Marcel Dekker, New York.
487. R. Gabasov and F. M. Kirillova (1977), Methods of optimal control, *J. Soviet Math.* **7**, 805–849.
488. R. Gabasov, F. M. Kirillova and B. S. Mordukhovich (1983), The ε-maximum principle for suboptimal controls, *Soviet Math. Dokl.* **27**, 95–99.
489. N. Gadhi (2006), Optimality conditions for a d. c. set-valued problem via the extremal principle, in *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, pp. 251–264. edited by S. Dempe and V. V. Kalashnikov, Springer, Berlin.
490. G. N. Galbraith (2001), Extended Hamilton-Jacobi characterization of value functions in optimal control, *SIAM J. Control Optim.* **39**, 281–305.
491. G. N. Galbraith (2002), Cosmically Lipschitz set-valued mappings, *Set-Valued Anal.* **10**, 331–360.
492. G. N. Galbraith (2004), Solution regularity in optimal control via subgradient analysis of the value function, *Set-Valued Anal.* **12**, 111–126.
493. G. N. Galbraith and R. B. Vinter (2003), Lipschitz continuity of optimal controls for state constrained problems, *SIAM J. Control Optim.* **42**, 1727–1744.
494. R. V. Gamkrelidze (1957), On the theory of optimal processes in linear systems, *Dokl. Akad. Nauk SSSR* **116**, 9–11.
495. R. V. Gamkrelidze (1962), On sliding optimal regimes, *Soviet Math. Dokl.* **3**, 559–561.
496. R. V. Gamkrelidze (1965), On some extremal problem in the theory of differential equations with applications to the theory of optimal control, *SIAM J. Control* **3**, 106–128.
497. R. V. Gamkrelidze (1978), *Principles of Optimal Control Theory*, Plenum Press, New York.
498. R. V. Gamkrelidze (1999), Discovery of the maximum principle, *J. Dynam. Control Syst.* **5**, 437–451.
499. J. Gauvin and F. Dubeau (1982), Differential properties of the marginal function in mathematical programming, *Math. Progr. Study* **19**, 101–119.

500. V. S. Gavrilov and M. I. Sumin (2004), Parametric optimization of nonlinear Goursat-Darboux systems with state constraints, *Comput. Maths. Math. Phys.* **44**, 949–968.
501. M. Geoffroy and M. Lassonde (2000), On a convergence of lower semicontinuous functions linked with the graph convergence of their subdifferentials, in *Constructive, Experimental and Nonlinear Analysis*, edited by M. Théra, Canad. Math. Soc. Conf. Proc. **27**, 93–109.
502. P. G. Georgiev and N. P. Zlateva (1996), Second-order subdifferential of $C^{1,1}$ functions and optimality conditions, *Set-Valued Anal.* **4**, 101–117.
503. W. Geremew, B. S. Mordukhovich and N. M. Nam (2005), Coderivative analysis of metric regularity for constraint and variational systems, preprint.
504. F. Giannessi (2005), *Constrained Optimization and Image Space Analysis*, published in two volumes Springer, Berlin.
505. J. R. Giles (1982), On the characterization of Asplund spaces, *J. Aust. Math. Soc.* **32**, 134–144.
506. B. Ginsburg and A. D. Ioffe (1996), The maximum principle in optimal control of systems governed by semilinear equations, in *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control*, edited by B. S. Mordukhovich and H. J. Sussmann, pp. 81–110, Springer, New York.
507. I. V. Girsanov (1972), *Lectures on Mathematical Theory of Extremum Problems*, Springer, Berlin.
508. B. M. Glover and B. D. Craven (1994), A Fritz John optimality condition using the approximate subdifferential, *J. Optim. Theory Appl.* **82**, 253–265.
509. B. M. Glover, B. D. Craven and S. D. Flåm (1993), A generalized Karush-Kuhn-Tucker optimality condition without constraint qualification using the approximate subdifferential, *Numer. Func. Anal. Optim.* **14**, 333–353.
510. B. M. Glover and D. Ralph (1994), First order approximations to nonsmooth mappings with applications to metric regularity, *Numer. Func. Anal. Optim.* **15**, 599–620.
511. R. Goebel (2004), Regularity of the optimal feedback and the value function in convex problems of optimal control, *Set-Valued Anal.* **12**, 127–145.
512. B. Gollan (1984), On the marginal function in nonlinear programming, *Math. Oper. Res.* **9**, 208–221.
513. S. B. Gorelik and B. S. Mordukhovich (1985), High-order necessary optimality conditions in neutral systems with some applications, *Izv. Akad Nauk BSSR. Ser. Fiz.-Mat. Nauk*, No. 6, 111–112; Depon. VINITI # 1154-85, Moscow.
514. S. B. Gorelik and B. S. Mordukhovich (1986), Legendre-Clebsch and Kelley conditions for nonlinear systems of neutral type with applications to existence of optimal control, *Izv. Akad Nauk BSSR. Ser. Fiz.-Mat. Nauk*, No. 65, 119; Depon. VINITI # 4057-85, Moscow.
515. V. V. Gorokhovik (1986), Epsilon-quasidifferentiability of real-valued functions and optimality conditions in extremal problems, *Math. Progr. Study* **29**, 203–218.
516. V. V. Gorokhovik (1990), *Convex and Nonsmooth Problems of Vector Optimization*, Nauka i Tekhnika, Minsk, Belarus.
517. V. V. Gorokhovik and P. P. Zabreiko (1998), Fréchet differentiability of multimappings, in *Nonlinear Analysis and Applications*, edited by I. V. Gaishun, pp. 34–49, Minsk, Belarus.

518. M. G. Govil and A. Mehra (2005), Epsilon-optimality for nonsmooth programming on a Hilbert space, in *Generalized Convexity, Generalized Monotonicity and Applications*, edited by A. Eberhard et al., Nonconvex Optimization and Its Applications, pp. 287–298, Springer, New York.
519. M. S. Gowda and R. Sznajder (1996), On the Lipshitzian properties of polyhedral multifunctions, *Math. Progr.* **74**, 267–278.
520. G. Grammel (2003), Towards fully discretized differential inclusions, *Set-Valued Anal.* **11**, 1–8.
521. K. A. Grasse (1985), A higher order sufficient condition for surjectivity, *Nonlinear Anal.* **9**, 87–96.
522. L. M. Graves (1950), Some mapping theorems, *Duke Math. J.* **17**, 111–114.
523. J. Guddat, F. Guerra Vasquez and H. T. Jongen (1990), *Parametric Optimization: Singularities, Pathfollowing and Jumps*, B. G. Tebner and John Wiley & Sons, Stuttgart and Chichester.
524. R. Guesnerie (1975), Pareto optimality in non-convex economies, *Econometrica* **43**, 1–29.
525. S. Guillaume (2000), Subdifferential evolution inclusions in nonconvex analysis, *Positivity* **4**, 357–395.
526. P. Gupta, S. Shiraishi and K. Yokoyama (2005), Epsilon-optimality without constraint qualifications for multiobjective fractional programs, *J. Nonlinear Convex Anal.* **6**, 347–357.
527. V. I. Gurman (1997), *The Extension Principle in Control Problems*, 2nd edition, Nauka, Moscow.
528. M. L. Gusakova (1972), Necessary optimality conditions for systems that are not solved with respect to the derivative, *Diff. Eq.* **8**, 1498–1500.
529. M. L. Gusakova (1974), On necessary optimality conditions for systems of neutral type, *Diff. Eq.* **810**, 1894–1897.
530. T. X. D. Ha (2003), The Ekeland variational principle for set-valued maps involving coderivatives, *J. Math. Anal. Appl.* **286**, 509–523.
531. T. X. D. Ha (2005), Some variants of the Ekeland variational principle for a set-valued map, *J. Optim. Theory Appl.* **124**, 187–206.
532. T. X. D. Ha (2005), Lagrange multipliers for set-valued problems associated with coderivatives, *J. Math. Anal. Appl.*, **311**, 647–663.
533. A. Habte (2005), *Applications of Variational Analysis to Welfare Economics*, Ph.D. dissertation, Department of Mathematics, Wayne State University, Detroit, Michigan.
534. N. Hadjisavvas, S. Komlosi and S. Schaible, eds. (2005), *Handbook of Generalized Convexity and Generalized Monotonicity*, Springer, Berlin.
535. W. W. Hager (1976), Rates of convergence for discrete approximations to unconstrained control problems, *SIAM J. Numer. Anal.* **13**, 1–14.
536. W. W. Hager (1979), Lipschitz continuity for constrained processes, *SIAM J. Control Optim.* **17**, 321–332.
537. A. Halanay (1968), Optimal control for systems with time lag, *SIAM J. Control* **6**, 215–234.
538. J. Hale (1977), *Theory of Functional Differential Equations*, Springer, New York.
539. H. Halkin (1964), On the necessary condition for the optimal control of nonlinear systems, *J. Analyse Math.* **12**, 1–82.

540. H. Halkin (1964), Optimal control for systems by nonlinear difference equations, in *Advances in Control Systems*, Vol. 1, pp. 173–196, Academic Press, New York.
541. H. Halkin (1970), A satisfactory treatment of equality and operator constraints in Dubovitskii-Milyutin optimization formalism, *J. Optim. Theory Appl.* **6**, 138–149.
542. H. Halkin (1972), Extremal properties of biconvex contingent equations, in *Ordinary Differential Equations*, edited by L. Weiss, pp. 109–119, Academic Press, New York.
543. H. Halkin (1974), Implicit functions and optimization problems without continuous differentiability of the data, *SIAM J. Control* **12**, 229–236.
544. H. Halkin (1976), Mathematical programming without differentiability, in *Calculus of Variations and Control Theory*, edited by D. L. Russel, pp. 279–287, Academic Press, New York.
545. H. Halkin (1978), Necessary conditions for optimal control problems with differentiable and nondifferentiable data, in *Mathematical Control Theory*, Lecture Notes Math. **680**, pp. 77–118, Springer, Berlin.
546. A. H. Hamel (2001), An epsilon-Lagrange multiplier rule for a mathematical programming problem on a Banach space, *Optimization* **49**, 137–150.
547. A. H. Hamel (2003), Phelps' lemma, Daneš' drop theorem and Ekeland's variational principle in locally convex spaces, *Proc. Amer. Math. Soc.* **131**, 3025–3038.
548. W. R. Hamilton (1834), On a general method employed in dynamics, by which the study of the motions of all free systems of attracting or repelling points is reduced to the search and differentiation of one central solution or characteristic function, *Philosophical Transactions of the Royal Society of London* **124**, 247–308.
549. W. L. Hare and A. Lewis (2005). Estimating tangent and normal cones, *Math. Oper. Res.*, **30**, 785–799.
550. P. T. Harker and J.-S. Pang (1990), Finite-dimensional variational inequalities and complementarity problems: a survey of theory, algorithms and applications, *Math. Progr.* **60**, 161–220.
551. J. Haslinger, M. Miettinen and P. D. Panagiotopoulos (1999), *Finite Elements Methods for Hemivariational Inequalities*, Kluwer, Dordrecht, The Netherlands.
552. F. Hausdorff (1927), *Mengenlehre*, Walter de Gruyter, Berlin.
553. R. Haydon (1990), A counterexample in several quastions about scattered compact spaces, *Bull. London Math. Soc.* **22**, 261–268.
554. Z.-X. He (1987), State constrained control problems governed by variational inequalities, *SIAM J. Control Optim.* **25**, 1119–1144.
555. R. Henrion (1995), Topological characterization of the approximate subdifferential in the finite-dimensional cases, *Math. Methods Oper. Res.* **41**, 161–173.
556. R. Henrion (1997), Topological properties of the approximate subdifferential, *J. Math. Anal. Appl.* **207**, 345–360.
557. R. Henrion (1997), Characterization of stability for cone increasing constraints in stochastic programming, *Set-Valued Anal.* **5**, 323–349.
558. R. Henrion (2004), Perturbation analysis of chance-constrained programs under variation of all constraint data, in *Dynamic Stochastic Optimization*, Lecture Notes Econ. Math. Syst. **523**, pp. 257–274, Springer, Berlin.

559. R. Henrion and A. Jourani (2002), Subdifferential conditions for calmness of convex constraints, *SIAM J. Optim.* **13**, 520–534.
560. R. Henrion, A. Jourani and J. V. Outrata (2002), On the calmness of a class of multifunctions, *SIAM J. Optim.* **13**, 603–618.
561. R. Henrion and J. V. Outrata (2001), A subdifferential condition for calmness of multifunctions, *J. Math. Anal. Appl.* **258**, 110–130.
562. R. Henrion and J. V. Outrata (2005), Calmness of constraint systems with applications, *Math. Progr.*, **104**, 437–464.
563. R. Henrion and W. Römisch (1999), Metric regularity and quantitative stability in stochastic programming with probabilistic constraints, *Math. Progr.* **84**, 55–88.
564. R. Henrion and W. Römisch (2000), Stability of solutions to chance constrained stochastic programs, in *Parametric Optimization and Related Topics V*, edited by J. Guddat et al., pp. 95–114, Peter Lang, Frankfurt, Germany.
565. M. R. Hestenes (1966), *Calculus of Variations and Optimal Control Theory*, Wiley, New York.
566. J. R. Hicks (1939), The foundations of welfare economics, *Econ. J.* **49**, 696–712.
567. D. Hilbert (1902), Mathematical problems, *Bull. Amer. Math. Soc.* **7**, 437–479.
568. W. Hildenbrand (1969), Pareto optimality for a measure space of economic agents, *Intern. Econ. Rev.* **10**, 363–372.
569. J.-B. Hiriart-Urruty (1979), New concepts in nondifferentiable programming, *Bull. Soc. Math. France* **60**, 57–85.
570. J.-B. Hiriart-Urruty (1979), Tangent cones, generalized gradients and mathematical programming in Banach spaces, *Math. Oper. Res.* **4**, 79–97.
571. J.-B. Hiriart-Urruty (1979), Refinements of necessary optimality conditions in nondifferentiable programming I, *Appl. Math. Optim.* **5**, 63–82.
572. J.-B. Hiriart-Urruty (1982), Refinements of necessary optimality conditions in nondifferentiable programming II, *Math. Progr. Study* **19**, 120–139.
573. J.-B. Hiriart-Urruty (1989), From convex optimization to nonconvex optimization: Necessary and sufficient conditions for global optimality, in *Nonconvex Optimization and Related Topics*, pp. 219–239, Plenum Press, New York.
574. J.-B. Hiriart-Urruty and Y. S. Ledyaev (1996), A note on the characterization of the global maxima of a (tangentially) convex function over a convex set, *J. Convex Anal.* **3**, 55–61.
575. J.-B. Hiriart-Urruty and C. Lemaréchal (1993), *Convex Analysis and Minimization Algorithms*, published in two volumes, Springer, Berlin.
576. J.-B. Hiriart-Urruty and P. Plazanet (1989), Moreau's decomposition theorem revisited, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **6**, 325–338.
577. J.-B. Hiriart-Urruty and A. Seeger (1989), Calculus rules on a new set-valued second-order derivative for convex functions, *Nonlinear Anal.* **13**, 721–738.
578. I. Hlaváček, J. Haslinger, J. Nečas and J. Lovišek (1988), *Solution of Variational Inequalities in Mechanics*, Springer, New York.
579. A. J. Hoffman (1952), On approximate solutions to systems of linear inequalities, *J. Nat. Bureau Stand.* **49**, 263–265.
580. R. B. Holmes (1975), *Geometric Functional Analysis and Its Applications*, Springer, New York.

581. L. HÖRMANDER (1954), Sur la function d'appui des ensdembles convexes dans un espace localement convexe, *Arkiv för Mat.* **3**, 181–186.
582. L. HÖRMANDER (1990), *The Analysis of Linear Partial Differential Operators I*, Springer, Berlin.
583. R. HORST, P. D. PARDALOS AND N. V. THOAI (2000), *Introduction to Global Optimization*, 2nd edition, Kluwer, Dordrecht, The Netherlands.
584. X. HU AND D. RALPH (2002), A note of sensitivity of value functions of mathematical programs with complementarity constraints, *Math. Progr.* **93**, 265–279.
585. X. HU, D. RALPH, E. K. RALPH, P. BARDSLEY AND M. C. FERRIS (2002), The effect of transmission capacities on competition in deregulated electricity markets, preprint.
586. D. K. HUGHES (1968), Variational and optimal control problems with delayed argument, *J. Optim. Theory Appl.* **2**, 1–14.
587. A. D. IOFFE (1979), Regular points of Lipschitz functions, *Trans. Amer. Math. Soc.* **251**, 61–69.
588. A. D. IOFFE (1979), Necessary and sufficient conditions for a local minimum, I: A reduction theorem and first order conditions, *SIAM J. Control Optim.* **17**, 245–250.
589. A. D. IOFFE (1981), Nonsmooth analysis: Differential calculus of non-differentiable mappings, *Trans. Amer. Math. Soc.* **266** (1981), 1–56.
590. A. D. IOFFE (1981), Sous-differentielles approaches de fonctions numériques, *C.R. Acad. Sci. Paris* **292**, 675–678.
591. A. D. IOFFE (1981), Calculus of Dini subdifferentials, CEREMADE Publication 8110, Universiteé de Paris IX "Dauphine".
592. A. D. IOFFE (1981), Approximate subdifferentials of nonconvex functions, CEREMADE Publication 8120, Universiteé de Paris IX "Dauphine".
593. A. D. IOFFE (1983), On subdifferentiability spaces, *Ann. New York Acad. Sci.* **410**, 107–119.
594. A. D. IOFFE (1984), Calculus of Dini subdifferentials and contingent coderivatives of set-valued maps, *Nonlinear Anal.* **8**, 517–539.
595. A. D. IOFFE (1984), Necessary conditions in nonsmooth optimization, *Math. Oper. Res.* **9**, 159–188.
596. A. D. IOFFE (1984), Approximate subdifferentials and applications, I: The finite dimensional theory, *Trans. Amer. Math. Soc.* **281**, 389–415.
597. A. D. IOFFE (1986), Approximate subdifferentials and applications, II: Functions on locally convex spaces, *Mathematika* **33**, 111–128.
598. A. D. IOFFE (1987), On the local surjection property, *Nonlinear Anal.* **11**, 565–592.
599. A. D. IOFFE (1989), Approximate subdifferentials and applications, III: The metric theory, *Mathematika* **36**, 1–38.
600. A. D. IOFFE (1990), Proximal analysis and approximate subdifferentials, *J. London Math. Soc.* **41**, 175–192.
601. A. D. IOFFE (1991), Variational analysis of composite functions: A formula for the lower second order epi-derivative, *J. Math. Anal. Appl.* **160**, 379–405.
602. A. D. IOFFE (1993), A Lagrange multiplier rule with small convex-valued subdifferentials for non-smooth problems of mathematical programming involving equality and non-functional constraints, *Math. Progr.* **72**, 137–145.

603. A. D. IOFFE (1996), Nonsmooth subdifferentials: Their calculus and applications, in *Proceedings of the First World Congress of Nonlinear Analysts*, edited by V. Lakshmikantham, pp. 2299–2310, De Gruyter, Berlin.

604. A. D. IOFFE (1997), Directional compactness and nonsmooth semi-Fredholm mappings, *Nonlinear Anal.* **29**, 201–219.

605. A. D. IOFFE (1997), Euler-Lagrange and Hamiltonian formalisms in dynamic optimization, *Trans. Amer. Amer. Soc.* **349**, 2871–2900.

606. A. D. IOFFE (1998), Fuzzy principles and characterization of trustworthiness, *Set-Valued Anal.* **6**, 265–276.

607. A. D. IOFFE (2000), Codirectional compactness, metric regularity and subdifferential calculus, in *Constructive, Experimental and Nonlinear Analysis*, edited by M. Théra, Canad. Math. Soc. Conf. Proc. **27**, 123–164.

608. A. D. IOFFE (2000), Metric regularity and subdifferential calculus, *Russian Math. Surveys* **55**, 501–558.

609. A. D. IOFFE (2003), On stability estimates for the regularity property of maps, in *Topological Methods, Variational Methods and Their Applications*, edited by H. Breziz et al., pp. 133–142, World Scientific Publishing, River Edge, New Jersey.

610. A. D. IOFFE (2003), On robustness for the regularity property of maps, *Control Cybernet.* **32**, 543–554.

611. A. D. IOFFE (2005), Optimality alternative: A non-variational approach to necessary conditions, in *Variational Analysis and Applications*, edited by F. Giannessi and A. Maugeri, pp. 531–552, Springer, Berlin.

612. A. D. IOFFE AND V. L. LEVIN (1972), Subdifferentials of convex functions, *Trudy Moscow Mat. Ob.* **26**, 3–73.

613. A. D. IOFFE AND T. MILOSZ (2002), On a characterization of $C^{1,1}$ functions, *Cybernetics and Systems Analysis*, No. 3, 3–13.

614. A. D. IOFFE AND J.-P. PENOT (1996), Subdifferentials of performance functions and calculus of coderivatives of set-valued mappings, *Serdica Math. J.* **22**, 359–384.

615. A. D. IOFFE AND J.-P. PENOT (1997), Limiting subhessians, limiting subjects and their calculus, *Trans. Amer. Math. Soc.* **349**, 789–807.

616. A. D. IOFFE AND R. T. ROCKAFELLAR (1996), The Euler and Weierstrass conditions for nonsmooth variational problems, *Calc. Var. Partial Diff. Eq.* **4**, 59–87.

617. A. D. IOFFE AND V. M. TIKHOMIROV (1968), Extensions of variational problems, *Trans. Moscow Math. Soc.* **18**, 207–273.

618. A. D. IOFFE AND V. M. TIKHOMIROV (1979), *Theory of Extremal Problems*, North-Holland, Amsterdam, The Netherlands.

619. A. D. IOFFE AND V. M. TIKHOMIROV (1997), Some remarks on variational principles, *Math. Notes* **61**, 248–253.

620. A. D. IOFFE AND A. J. ZASLAVSKI (2000), Variational principles and well-posedness in optimization and calculus of variations, *SIAM J. Control Optim.* **38**, 566–581.

621. A. N. IUSEM (1998), On some properties of generalized proximal point methods for variational inequalities, *J. Optim. Theory Appl.* **96**, 337–362.

622. M. IVANOV (2004), Sequential representation formulae for G-subdifferential and Clarke subdifferential in smooth Banach spaces, *J. Convex Anal.* **11**, 179–196.

623. A. F. IZMAILOV (2006), *Sensitivity in Optimization*, Fizmatlit, Moscow.
624. A. F. IZMAILOV AND M. V. SOLODOV (2001), Error bounds for 2-regular mappings with Lipschitzian derivatives and its applications, *Math. Progr.* **89**, 413–435.
625. C. G. J. JACOBI (1866), *Volesungen über Dynamik*, edited by A. Clebsch, Verlag Georg Reimer, Berlin.
626. M. Q. JACOBS AND C. E. LANGENHOP (1976), Criteria for function space controllability of linear neutral systems, *SIAM J. Control Optim.* **14**, 1009–1048.
627. J. JAHN (2004), *Vector Optimization: Theory, Applications and Extensions*, Springer, Berlin.
628. J. JAHN, A. A. KHAN AND P. ZEILINGER (2005), Second order optimality conditions in set-valued optimization, *J. Optim. Theory Appl.* **125**, 331–347.
629. R. JANIN (1973), Sur une classe de fonctions sous-linéarisables, *C. R. Acad. Sci. Paris* **277**, 265–267.
630. V. JEYAKUMAR AND D. T. LUC (1998), Approximate Jacobian matrices for nonsmooth continuous maps and C^1-optimization, *SIAM J. Control Optim.* **36**, 1815–1832.
631. V. JEYAKUMAR AND N. D. YEN (2004), Solution stability, regularity and implicit functions for nonsmooth continuous systems, *SIAM J. Optim.* **14**, 1106–1127.
632. H. JIANG AND D. RALPH (2000), Smooth SQP methods for mathematical programs with nonlinear complementarity constraints, *SIAM J. Optim.* **10**, 779–808.
633. A. JOFRÉ (2000), A second-welfare theorem in nonconvex economics, in *Constructive, Experimental and Nonlinear Analysis*, edited by M. Théra, Canad. Math. Soc. Conf. Proc. **27**, 175–184.
634. A. JOFRÉ, D. T. LUC AND M. THÉRA (1998), Epsilon-sundifferentials and epsilon-monotonicity, *Nonlinear Anal.* **33**, 71–90.
635. A. JOFRÉ AND J. RIVERA C. (2006), A nonconvex separation property and some applications, *Math. Progr.*, **108**, 37–51.
636. A. JOFRÉ, R. T. ROCKAFELLAR AND R. J-B. WETS (2005), A variational inequality scheme for determining an economic equilibrium of classical or extended type, in *Variational Analysis and Applications*, edited by F. Giannessi and A. Maugeri, pp. 553–578, Springer, Berlin.
637. A. JOFRÉ, R. T. ROCKAFELLAR AND R. J-B. WETS (2007), Variational inequalities and economic equilibrium, *Math. Oper. Res.*, **32**, 32–50.
638. F. JOHN (1948), Extremum problems with inequalities as side conditions, in *Studies and Essays: Courant Anniversary Volume*, edited by K. O. Fredrichs, O. E. Neugebauer and J. J. Stoker, pp. 187–204, Wiley, New York.
639. H. T. JONGEN, D. KLATTE ANr K. TAMMER (1990), Implicit functions and sensitivity of stationary points, *Math. Progr.* **49**, 123–138.
640. H. T. JONGEN, T. MÖBERT, J.-J. RÜCKMANN AND K. TAMMER (1987), Implicit functions and sensitivity analysis of stationary points, *Linear Algerbra Appl.* **95**, 97–109.
641. H. T. JONGEN, J.-J. RÜCKMANN AND G.-W. WEBER (1994), One-parametric semi-infinite optimization: On the stability of the feasible set, *SIAM J. Optim.* **4**, 637–648.
642. E. JOUINI (1988), A remark on Clarke's normal cone and the marginal cost pricing rule, *J. Math. Econ.* **17**, 309–315.

643. A. JOURANI (1995), Compactly epi-Lipschitzian sets and A-subdifferentials in WT-spaces, *Optimization* **34**, 1–17.
644. A. JOURANI (1995), Intersection formulae and the marginal function in Banach spaces, *J. Math. Anal. Appl.* **192**, 867–891.
645. A. JOURANI (1998), Necessary conditions for extremality and separation theorems with applications to multobjective optimization, *Optimization* **44**, 327–350.
646. A. JOURANI (1999), Limit superior of subdifferentials of uniformly convergent functions, *Positivity* **3**, 33–47.
647. A. JOURANI (2000), Hoffman's error bound, local controllability and sensitivity analysis, *SIAM J. Control Optim.* **38**, 947–970.
648. A. JOURANI (2003), On a class of compactly epi-Lipschitzian sets, *Nonlinear Anal.* **54**, 471–483.
649. A. JOURANI (2006), Weak regularity of functions and sets in Asplund spaces, *Nonlinear Anal.*, **65**, 660–676.
650. A. JOURANI AND M. THÉRA (1998), On the limiting Fréchet ε-sibdifferentials, in *Generalized Convexity, Generalized Monotonicity: Recent Results*, edited by J. P. Crouzeix et al., Nonconvex Optim. Appl. **27**, pp. 185–198, Kluwer, Dordrecht, The Netherlands.
651. A. JOURANI AND L. THIBAULT (1990), Approximate subdifferential and metric regularity: The finite-dimensional case, *Math. Progr.* **47**, 203–218.
652. A. JOURANI AND L. THIBAULT (1990), The use of metric graphical regularity in approximate subdifferential calculus rules in finite dimensions, *Optimization* **21**, 1–11.
653. A. JOURANI AND L. THIBAULT (1993), The approximate subdifferential of composite functions, *Bull. Austral. Math. Soc.* **47**, 443–445.
654. A. JOURANI AND L. THIBAULT (1994), A note on Fréchet and approximate subdifferentials of composite functions, *Bull. Austral. Math. Soc.* **49**, 111–116.
655. A. JOURANI AND L. THIBAULT (1995), Verifiable conditions for openness, metric regularity of multivalued mappings, *Trans. Amer. Math. Soc.* **347**, 1225–1268.
656. A. JOURANI AND L. THIBAULT (1995), Metric regularity for strongly compactly Lipschitzian mappings, *Nonlinear Anal.* **24**, 229–240.
657. A. JOURANI AND L. THIBAULT (1996), Metric regularity and subdifferential calculus in Banach spaces, *Set-Valued Anal.* **3**, 87–100.
658. A. JOURANI AND L. THIBAULT (1996), Extensions of subdifferential calculus rules in Banach spaces, *Canad. J. Math.* **48**, 834–848.
659. A. JOURANI AND L. THIBAULT (1998), Chain rules for coderivatives of multivalued mappings in Banach spaces, *Proc. Amer. Math. Soc.* **126**, 1479–1485.
660. A. JOURANI AND L. THIBAULT (1998), Qualification conditions for calculus rules of coderivatives of multivalued mappings, *J. Math. Anal. Appl.* **218**, 66–81.
661. A. JOURANI AND L. THIBAULT (1999), Coderivatives of multivalued mappings, locally compact cones and metric regularity, *Nonlinear Anal.* **35**, 925–945.
662. A. JOURANI AND J. J. YE (2005), Error bounds for eigenvalue and semidefinite matrix inequality systems, *Math. Progr.*, **104**, 525–540.
663. G. A. KAMENSKII AND E. A. HVILON (1968), Optimality conditions for systems with deviating argument, in *Trudy Sem. Teor. Diff. Urav. Otklon. Argum.* **6**, 213–222.

664. L. V. KANTOROVICH (1940), On an efficient method for solving some classes of extremum problems, *Dokl. Akad. Nauk SSSR* **28**, 212–215.
665. W. KARUSH (1939), *Minima of Functions of Several Variables with Inequalities as Side Conditions*, Master Thesis, Department of Mathematics, University of Chicago.
666. B. KAŚKOSZ AND S. LOJASIEWICZ, JR. (1985), A maximum principle for generalized control systems, *Nonlinear Anal.* **9**, 109–130.
667. B. KAŚKOSZ AND S. LOJASIEWICZ, JR. (1992), Lagrange-type extremal trajectories in differential inclusions, *Syst. Control Lett.* **19**, 241–247.
668. G. A. KENT (1971), A maximum principle for optimal control problems with neutral functional differential systems, *Bull. Amer. Math. Soc.* **77**, 565–570.
669. M. A. KHAN (1988) Ioffe's normal cone and the foundations of welfare economics: An example, *Econ. Lett.* **28**, 5–19.
670. M. A. KHAN (1991), Ioffe's normal cone and the foundations of welfare economics: The infinite dimensional theory, *J. Math. Anal. Appl.* **161**, 284–298.
671. M. A. KHAN (1999), The Mordukhovich normal cone and the foundations of welfare economics, *J. Public Econ. Theory* **1**, 309–338.
672. M. A. KHAN AND S. RASHID (1975), Nonconvexity and Pareto optimality in large markets, *Inter. Econ. Rev.* **16**, 222–245.
673. M. A. KHAN AND R. VOHRA (1987), An extension of the second welfare theorem to economies with non-convexities and public goods, *Quartery J. Econ.* **102**, 223–245.
674. M. A. KHAN AND R. VOHRA (1988), Pareto optimal allocations of nonconvex economies in locally convex spaces, *Nonlinear Anal.* **12**, 943–950.
675. M. A. KHAN AND R. VOHRA (1988), On approximate decentralization of Pareto optimal allocations in locally convex spaces, *J. Approx. Theory* **52**, 149–161.
676. P. Q. KHANH (1986), An induction theorem and general open mapping theorem, *J. Math. Anal. Appl.* **118**, 519–536.
677. P. Q. KHANH (1989), On general open mapping theorems, *J. Math. Anal. Appl.* **144**, 305–312.
678. G. L. KHARATISHVILI (1961), The maximum principle in the theory of optimal processes with a delay, *Soviet Math. Dokl.* **2**, 28–32.
679. G. L. KHARATISHVILI AND T. A. TADUMADZE (1998), A nonlinear optimal control problem with variable delays, nonfixed initial moment, and piecewise-continuous prehistory, *Proc. Steklov Inst. Math.* **220**, 233–252.
680. D. KINDERLEHRER AND G. STAMPACCHIA (1980), *An Introduction to Variational Inequalities and Their Applications*, Academic Press, New York.
681. A. J. KING AND R. T. ROCKAFELLAR (1992), Sensitivity analysis for nonsmooth generalized equations, *Math. Progr.* **55**, 193–212.
682. M. KISIELEWICZ (1991), *Differential Inclusions and Optimal Control*, Kluwer, Dordrecht, The Netherlands.
683. K. C. KIWIEL (1985), *Methods of Descent for Nondifferentiable Optimization*, Springer, Berlin.
684. D. KLATTE (1994), On quantitative stability for non-isolated minima, *Control Cybernet.* **23**, 183–200.
685. D. KLATTE AND R. HENRION (1998), Regularity and stability in nonlinear semi-infinite optimization, in *Semi-Infinite Programming*, edited by R. Reemtsen and J.-J. Rückmann, pp. 69–102, Kluwer, Dordrecht, The Netherlands.

686. D. Klatte and B. Kummer (2002), *Nonsmooth Equations in Optimization: Regularity, Calculus, Methods, and Applications*, Kluwer, Boston, Massachusetts.
687. D. Klatte and B. Kummer (2002), Constrained minima and Lipschitzian penalties in metric spaces, *SIAM J. Optim.* **13**, 619–633.
688. M. Kocan and A. Świech (1996), Perturbed optimization on product spaces, *Nonlinear Anal.* **26**, 81–90.
689. M. Kočvara, M. Kružik and J. V. Outrata (2005), On the control of an evolutionary equilibrium in micromagnetics, in *Optimization with Multivalued Mappings: Theory, Applications and Algorithms*, edited by S. Dempe and V. V. Kalashnikov, Springer, Berlin.
690. M. Kočvara and J. V. Outrata (2004), Optimization problems with equilibrium constraints and their numerical solutions, *Math. Progr.* **101**, 119–149.
691. M. Kočvara and J. V. Outrata (2005), On the modeling and control of delamination processes, in *Control and Boundary Analysis*, edited by J. Cagnol and J.-P. Zolésio, Lecture Notes Pure Applied Math. **240**, pp. 169–185, Marcel Dekker, New York.
692. M. Kojima (1980), Strongly stable stationary solutions in nonlinear programs, in *Analysis and Computation of Fixed Points*, edited by S. M. Robinson, pp. 93–138, Academic Press, New York.
693. P. V. Kokotović, H. K. Khalil and J. O'Reilly (1986), *Singular Perturbations in Control Analysis and Design*, Academic Press, New York.
694. V. B. Kolmanovskii and A. D. Myshkis (1992), *Applied Theory of Functional Differential Equations*, Kluwer, Dordrecht, The Netherlands.
695. V. B. Kolmanovskii and V. R. Nosov (1984), Neutral-type systems with aftereffects, *Autom. Remote Control* **45**, 1–28.
696. V. B. Kolmanovskii and L. E. Shaikhet (1996), *Control of Systems with Aftereffect*, Academic Press, New York.
697. S. Komlosi, T. Rapcsák and S. Schaible, eds. (1994), *Generalized Convexity*, Lecture Notes Econ. Math. Syst. **405**, Springer, Berlin.
698. E. Kostina and O. Kostyukova (2006), Generalized implicit function theorem and its applications to parametric optimal control, *J. Math. Anal. Appl.*, **320**, 736–755.
699. S. G. Krantz and H. R. Parks (2002), *The Implicit Function Theorem: History, Theory, and Applications*, Birkhäuser, Boston, Massachusetts.
700. N. N. Krasovskii (1964), Optimal processes in systems with time lag, in *Automatics and Remote Control Theory*, edited by L. Butterths, pp. 327–332, Oldenburg, Munish, Germany.
701. N. N. Krasovskii and N. Y. Lukoyanov (2000), Equations of Hamilton-Jacobi type in hereditary systems: Minimax solutions, *Proc. Steklov Inst. Math.: Control in Dynamic Systems, suppl. 1*, 136–153.
702. N. N. Krasovskii and A. I. Subbotin (1988), *Game-Theoretical Control Problems*, Springer, New York.
703. A. J. Krener (1977), The high order maximum principle and its applications to singular extremals, *SIAM J. Control Optim.* **15**, 256–293.
704. V. F. Krotov (1996), *Global Methods in Optimal Control*, Marcel Dekker, New York.
705. A. Y. Kruger (1981), Epsilon-semidifferentials and epsilon-normal elements, Depon. VINITI #1331-81, Moscow.

706. A. Y. KRUGER (1981), *Generalized Differentials of Nonsmooth Functions and Necessary Conditions for an Extremum*, Ph.D. dissertation, Department of Mathematics, Belarus State University, Minsk, Belarus.
707. A. Y. KRUGER (1985), Generalized differentials of nonsmooth functions and necessary conditions for an extremum, *Siberian Math. J.* **26**, 370–379.
708. A. Y. KRUGER (1985), Properties of generalized differentials, *Siberian Math. J.* **26**, 822–832.
709. A. Y. KRUGER (1988), A covering theorem for set-valued mappings, *Optimization* **19**, 763–780.
710. A. Y. KRUGER (1997), Strict ε-semidifferentials and extremality conditions, *Dokl. Nat. Akad. Belarus* **41**, 21–26.
711. A. Y. KRUGER (1998), On extremality of set systems, *Dokl. Nat. Akad. Belarus* **42**, 24–28.
712. A. Y. KRUGER (2002), Strict (ε, δ)-semidifferentials and extremality conditions, *Optimization* **51**, 539–554.
713. A. Y. KRUGER (2003), On Fréchet subdifferentials, *J. Math. Sci.* **116**, 3325–3358.
714. A. Y. KRUGER (2004), Weak stationarity: Eliminating the gap between necessary and sufficient conditions, *Optimization* **53**, 147–164.
715. A. Y. KRUGER (2005), Stationarity and regularity of set systems, *Pacific J. Optim.* **1**, 101–126.
716. A. Y. KRUGER (2006), About regularity of set systems, *Set-Valued Anal.* **14**, 187–206.
717. A. Y. KRUGER AND B. S. MORDUKHOVICH (1978), Minimization of nonsmooth functionals in optimal control problems, *Eng. Cybernetics* **16**, 126–133.
718. A. Y. KRUGER AND B. S. MORDUKHOVICH (1980), Extremal points and the Euler equation in nonsmooth optimization, *Dokl. Akad. Nauk BSSR* **24**, 684–687.
719. A. Y. KRUGER AND B. S. MORDUKHOVICH (1980), Generalized normals and derivatives, and necessary optimality conditions in nondifferential programming, Parts I and II, Depon. VINITI: I# 408-80, II# 494-80, Moscow.
720. S. N. KRUZHKOV (1960), The Cauchy problem in the large for certain nonlinear first order differential equations, *Soviet Math. Dokl.* **1**, 474–477.
721. A. V. KRYAZHIMSKII AND Y. S. OSIPOV (1995), On evolutionary-differential games, *Proc. Steklov Math. Inst.* **211**, 257–287.
722. E. I. KUGUSHEV (1973), Maximum principle in optimal control systems with nonsmooth right-hand side, *Vest. Moskov. Univ. Ser. Mat. Mekh.*, No. 3, 107–113.
723. H. W. KUHN (1976), Nonlinear programming: A historical survey, in *Nonlinear Programming*, edited by R. W. Cottle and C. E. Lemke, pp. 1–26, American Mathematical Society, Providence, Rhode Island.
724. H. W. KUHN AND A. W. TUCKER (1951), Nonlinear programming, in *Proceedings of the Second Berkley Symposium on Mathematical Statistics and Probability*, edited by J. Neyman, pp. 481–492, University of California Press, Berkeley, California.
725. B. KUMMER (1991), Lipschitzian inverse functions, directional derivatives and applications in $C^{1,1}$ optimization, *J. Optim. Theory Appl.* **70**, 561–582.
726. B. KUMMER (1991), An implicit function theorem for $C^{0,1}$-equations and parametric $C^{1,1}$-optimization, *J. Math. Anal. Appl.* **158**, 35–46.

727. B. KUMMER (1999), Metric regularity: Characterizations, nonsmooth variations and successive approximation, *Optimization* **46**, 247–281.
728. B. KUMMER (2000), Inverse functions of pseudo regular mappings and regularity conditions, *Math. Progr.* **88**, 313–339.
729. L. KUNTZ AND S. SCHOLTES (1994), Structural analysis of nonsmooth mappings, inverse functions, and metric projections, *J. Math. Anal. Appl.* **188**, 346–386.
730. G. A. KURINA (1979), Application of the method of tents to an optimal control problem for a differential equation with a singular matrix multiplying the derivative, *Diff. Eq.* **15**, 417–423.
731. A. B. KURZHANSKII (1977), *Control and Observation under Conditions of Uncertainty*, Nauka, Moscow.
732. A. B. KURZHANSKII AND I. VÁLYI (1997), *Ellipsoidal Calculus for Estimation and Control*, Burlhäuser, Boston, Massachusetts.
733. A. G. KUSRAEV AND S. S. KUTATELADZE (1995), *Subdifferentials: Theory and Applications*, Kluwer, Dordrecht, The Netherlands.
734. J. KYPARISIS (1992), Parametric variational inequalities with multivalued solution sets, *Math. Oper. Res.* **17**, 341–364.
735. O. A. LADYZHENSKAYA, V. A. SOLONNIKOV AND N. N. URALZEVA (1968), *Linear and Quasilinear Equations of Parabolic Type*, American Mathematical Society, Providence, Rhode Island.
736. J. E. LAGNESE (1989), *Boundary Stabilization of Thin Plates*, SIAM Publications, Philadelphia, Pennsylvania.
737. J. L. LAGRANGE (1813), *Thérie des fonctions analytique*, Paris.
738. O. LANGE (1942), The foundations of welfare economics, *Econometrica* **10**, 215–228.
739. I. LASIECKA (1980), Unified theory of abstract parabolic boundary problems – a semigroup approach, *Appl. Math. Optim.* **6**, 287–333.
740. I. LASIECKA, J.-L. LIONS AND R. TRIGGIANI (1986), Nonhomogeneous boundary value problems for second-order hyperbolic operators, *J. Mat. Pures Appl.* **65**, 149–192.
741. I. LASIECKA AND J. SOKOLOWSKI (1988), Regularity and strong convergence of a variational approximation to a nonhomogeneous Dirichlet hyperbolic boundary problem, *SIAM J. Math. Anal.* **19**, 528–540.
742. I. LASIECKA AND R. TRIGGIANI (1983), Dirichlet boundary control problems for parabolic equations with quadratic cost: Analyticity and Riccati feedback synthesis, *SIAM J. Control Optim.* **21**, 41–67.
743. I. LASIECKA AND R. TRIGGIANI (1987), The regulator problem for parabolic equations with Dirichlet boundary control, I: Riccati's feedback synthesis and regularity of optimal solutions, *Appl. Math. Optim.* **16**, 147–168.
744. I. LASIECKA AND R. TRIGGIANI (1990), Sharp regularity theory for second order hyperbolic equations of Neumann type, I: L_2 nonhomogeneous data, *Ann. Mat. Pura Appl.* **157**, 285–367.
745. I. LASIECKA AND R. TRIGGIANI (1991), Regularity theory of hyperbolic equations with non-homogeneous Neumann boundary conditions, II: General boundary data, *J. Diff. Eq.* **94**, 112–164.
746. I. LASIECKA AND R. TRIGGIANI (2000), *Control Theory for Partial Differential Equations: Continuous and Approximation Theory*, published in two volumes, Cambridge University Press, Cambridge, UK.

747. M. LASSONDE (2001), First-order rules for nonsmooth constrained optimization, *Nonlinear Anal.* **44**, 1031–1056.
748. E. B. LEACH (1961), A note on inverse function theorem, *Proc. Amer. Math. Soc.* **12**, 694–697.
749. G. LEBOURG (1975), Valeur moyenne pour gradient généraliseé, *C. R. Acad. Sci. Paris* **281**, 795–798.
750. Y. S. LEDYAEV (2004), On generic existence and uniqueness in nonconvex optimal control, *Set-Valued Anal.* **12**, 147–162.
751. Y. S. LEDYAEV AND Q. J. ZHU (1999), Implicit multifunction theorem, *Set-Valued Anal.* **7**, 209–238.
752. Y. S. LEDYAEV AND Q. J. ZHU (2004), Techniques for nonsmooth analysis on smooth manifolds I: Techniques for local problems, in *Optimal Control, Stabilization and Nonsmooth Analysis*, edited by M. de Queiroz, M. Malisoff and P. Wolenski, Lecture Notes Cont. Inf. Sci. **301**, pp. 283–287, Springer, New York.
753. Y. S. LEDYAEV AND Q. J. ZHU (2004), Techniques for nonsmooth analysis on smooth manifolds II: Using deformations and flowes, in *Optimal Control, Stabilization and Nonsmooth Analysis*, edited by M. de Queiroz, M. Malisoff and P. Wolenski, Lecture Notes Cont. Inf. Sci. **301**, pp. 299–311, Springer, New York.
754. Y. S. LEDYAEV AND Q. J. ZHU (2005), Nonsmooth analysis on smooth manifolds, *Trans. Amer. Math. Soc.*, to appear.
755. G. M. LEE, N. N. TAM AND N. D. YEN (2005), Normal coderivative for multifunctions and implicit function theorems, preprint.
756. U. LEDZEWICZ AND H. SCHÄTTLER (1998), High-order approximations and generalized necessary conditions for optimality, *SIAM J. Control Optim.* **37**, 33–56.
757. G. W. LEIBNIZ (1696), Letter to Johann Bernoulli of July 31, 1696, Hanover, Germany.
758. G. LEITMANN (1978), On generalized Stackelberg strategies, *J. Optim. Theory Appl.* **59**, 637–643.
759. C. LEMARÉCHAL, F. OUSTRY AND C. SAGASTIZÁBAL (2000), The $\mathcal{U}$-Lagrangian of a convex function, *Trans. Amer. Math. Soc.* **352**, 711–729.
760. F. LEMPIO (1995), Euler's method revisited, *Proc. Steklov Inst. Math.* **211**, 473–494.
761. F. LEMPIO AND V. M. VELIOV (1998), Discrete approximations to differential inclusions, *Mitteilungen der GAMM* **21**, 101–135.
762. S. LENHART, V. PROTOPESCU AND S. STOJANOVIĆ (1993), A minimax problem for semilinear nonlocal competitive systems, *Appl. Math. Optim.* **28**, 113–132.
763. E. S. LEVITIN (1994), *Perturbation Theory in Mathematical Programming and Its Applications*, Wiley, New York.
764. E. S. LEVITIN, A. A. MILYUTIN AND N. P. OSMOLOVSKII (1978), Higher order conditions for a local minimum in the problems with constraints, *Russian Math. Surveys* **33**, 97–168.
765. A. B. LEVY (1993), Second-order epi-derivatives of integral functionals, *Set-Valued Anal.* **1**, 379–392.
766. A. B. LEVY (1996), Implicit multifunction theorems for the sensitivity analysis of variational conditions, *Math. Progr.* **74**, 333–350.

767. A. B. LEVY (2001), Lipschitzian multifunctions and a Lipschitzian inverse function theorem, *Math. Oper. Res.* **26**, 105–118.
768. A. B. LEVY (2001), Solution stability from general principles, *SIAM J. Control Optim.* **40**, 209–238.
769. A. B. LEVY AND B. S. MORDUKHOVICH (2004), Coderivatives in parametric optimization, *Math. Progr.* **99**, 311–327.
770. A. B. LEVY AND R. A. POLIQUIN (1997), Characterizing the single-valuedness of multifunctions, *Set-Valued Anal.* **5**, 351–364.
771. A. B. LEVY, R. A. POLIQUIN AND R. T. ROCKAFELLAR (2000), Stability of locally optimal solutions, *SIAM J. Optim.* **10**, 580–604.
772. A. B. LEVY, R. A. POLIQUIN AND L. THIBAULT (1995), A partial extension of Attouch's theorem and its applications to second-order epi-differentiation, *Trans. Amer. Math. Soc.* **347**, 1269–1294.
773. A. B. LEVY AND R. T. ROCKAFELLAR (1994), Sensitivity analysis of solutions to generalized equations, *Trans. Amer. Math. Soc.* **345**, 661–671.
774. A. B. LEVY AND R. T. ROCKAFELLAR (1996), Variational conditions and the proto-differentiation of partial subgradient mappings, *Nonlinear Anal.* **26**, 1951–1964.
775. A. S. LEWIS (1999), Nonsmooth analysis of eigenvalues, *Math. Progr.* **84**, 1–24.
776. A. S. LEWIS (1999), Ill-conditioned convex processes and conic linear systems, *Math. Oper. Res.* **23**, 829–834.
777. A. S. LEWIS (2001), Ill-conditioned inclusions, *Set-Valued Anal.* **9**, 375–381.
778. A. S. LEWIS (2003), Active sets, nonsmoothness and sensitivity, *SIAM J. Optim.* **13**, 702–725.
779. A. S. LEWIS (2003), The mathematics of eigenvalue optimization, *Math. Progr.* **97**, 155–176.
780. A. S. LEWIS (2004), The structured distance to ill-posedness for conic systems, *Math. Oper. Res.* **29**, 776–785.
781. A. S. LEWIS AND M. L. OVERTON (1996), Eigenvalue optimization, *Acta Numerica* **5**, 149–190.
782. A. S. LEWIS AND H. S. SENDOV (2005), Nonsmooth analysis of singular values. Part I: Theory, *Set-Valued Anal.* **13**, 213–241.
783. A. S. LEWIS AND H. S. SENDOV (2005), Nonsmooth analysis of singular values. Part II: Applications, *Set-Valued Anal.* **13**, 243–264.
784. W. LI AND I. SINGER (1998), Global error bounds for convex multifunctions, *Math. Oper. Res.* **23**, 443–462.
785. Y. LI AND S. SHI (2000), A generalization of Ekeland's ε-variational principle and its Borwein-Preiss variant, *J. Math. Anal. Appl.* **246**, 308–319.
786. X. LI AND Y. YAO (1981), On optimal control for distributed parameter systems, in *Proc. IFAC 8th Triennial World Congress*, pp. 207–212, Kyoto, Japan.
787. X. LI AND Y. YAO (1985), Maximum principle of distributed parameter systems with time lags, in *Lecture Notes Cont. Inform. Sci.* **75**, pp. 410–427, Springer, Berlin.
788. X. LI AND J. YONG (1991), Necessary conditions of optimal control for distributed parameter systems, *SIAM J. Control Optim.* **29**, 895–908.
789. X. LI AND J. YONG (1995), *Optimal Control Theory for Infinite-Dimensional Systems*, Birkhäuser, Boston, Massachusetts.

790. L.-J. LIN AND W.-S. DU (2006), Ekeland's variational principles and the existence of nonconvex equilibria and minimax inequalities in complete metric spaces, *J. Math. Anal. Appl.*, **323**, 360–370.
791. J.-L. LIONS (1971), *Optimal Control of Systems Governed by Partial Differential Equations*, Springer, Berlin.
792. J.-L. LIONS (1983), *Contrôle des Systémes Distribués Singuliers*, Gauthier-Villars, Paris.
793. P.-L. LIONS (1982), *Generalized Solutions of Hamilton-Jacobi Equations*, Pitman, Boston, Massachusetts.
794. J.-L. LIONS AND E. MAGENES (1972), *Nonhomogeneous Boundary Value Problems and Applications*, Springer, New York.
795. J.-L. LIONS AND G. STAMPACCHIA (1967), Variational inequalities, *Comm. Pure Appl. Math.* , 493–519.
796. R. LIPSCHITZ (1877), *Lehrbuch der Analysis*, Cohen & Sohn, Bonn.
797. J. M. LIU (1995), Strong stability in variational inequalities, *SIAM J. Control Optim.* **33**, 725–749.
798. P. D. LOEWEN (1987), The proximal normal formula in Hilbert spaces, *Nonlinear Anal.* **11**, 979–995.
799. P. D. LOEWEN (1988), The proximal subgradient formula in Banach spaces, *Canad. Math. Bull.* **31**, 353–361.
800. P. D. LOEWEN (1992), Limits of Fréchet normals in nonsmooth analysis, in *Optimization and Nonlinear Analysis*, edited by A. Ioffe, L. Marcus and S. Reich, Pitman Research Notes Math. Ser. **244**, pp. 178–188, Longman, Harlow, Essex, UK.
801. P. D. LOEWEN, *Optimal Control via Nonsmooth Analysis* (1993), American Mathematical Society, Providence, Rhode Island.
802. P. D. LOEWEN (1994), A mean value theorem for Fréchet subgradients, *Nonlinear Anal.* **23**, 1365–1381.
803. P. D. LOEWEN, F. H. CLARKE AND R. B. VINTER (1988), Differential inclusions with free time, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **5**, 573–593.
804. P. D. LOEWEN AND R. T. ROCKAFELLAR (1991), The adjoint arcs in nonsmooth optimization, *Trans. Amer. Math. Soc.* **325**, 39–72.
805. P. D. LOEWEN AND R. T. ROCKAFELLAR (1994), Optimal control of unbounded differential inclusions, *SIAM J. Control Optim.* **32**, 442–470.
806. P. D. LOEWEN AND R. T. ROCKAFELLAR (1996), New necessary conditions for the generalized problem of Bolza, *SIAM J. Control Optim.* **34**, 1496–1511.
807. P. D. LOEWEN AND R. T. ROCKAFELLAR (1997), Bolza problem with general time constraints, *SIAM J. Control Optim.* **35**, 2050–2069.
808. P. D. LOEWEN AND R. B. VINTER (1987), Pontryagin type necessary conditions for differential inclusions, *Syst. Control Lett.* **9**, 263–265.
809. P. D. LOEWEN AND X. WANG (2001), A generalized variational principle, *Canad. J. Math.* **53**, 1174–1193.
810. S. LOJASIEWICZ, JR. (1992), Local controllability of parameterized differential equations, preprint.
811. P. LORIDAN (1982), Necessary conditions for ε-optimality, *Math. Progr. Study* **19**, 140–152.
812. P. LORIDAN AND J. MORGAN (1989), New results on approximate solutions in two level optimization, *Optimization* **20**, 819–836.

813. D. T. LUC (1989), *Theory of Vector Optimization*, Springer, Berlin.
814. D. T. LUC (1996), A strong mean value theorem and applications, *Nonlinear Anal.* **26**, 915–923.
815. D. T. LUC (2005), Chain rules for approximate Jacobians of continuous functions, *Nonlinear Anal.* **61**, 97–114.
816. Y. LUCET AND J. J. YE (2001), Sensitivity analysis of the value function for optimization problems with variational inequality constraints, *SIAM J. Optim.* **40**, 699–723.
817. D. R. LUKE, J. V. BURKE AND R. G. LYON (2002), Optical wavefront reconstruction: Theory and numerical methods, *SIAM Rev.* **44**, 169–224.
818. Y. LUO AND A. EBERHARD (2005), Comparizon principles for viscosity solutions of elliptic equations via fuzzy sum rule, *J. Math. Anal. Appl.* **307**, 736–752.
819. Y. LUO AND A. EBERHARD (2005), An application of $C^{1,1}$ approximation to comparison principles for viscosity solutions of curvature equations, preprint.
820. Z. Q. LUO, J.-S. PANG AND D. RALPH (1996), *Mathematical Programs with Equilibrium Constraints*, Cambridge University Press, Cambridge, UK.
821. K. A. LURIE (1993), *Applied Optimal Control of Distributed Systems*, Plenum Press, New York.
822. A. A. LYAPUNOV (1940), Sur les founctions-vecteurs complétement additives, *Izvest. Akad. Nauk SSSR, Ser. Mat.* **3**, 465–478.
823. S. I. LYASHKO (2002), *Generalized Optimal Control of Linear Systems with Distributed Parameters*, Kluwer, Dordrecht, The Netherlands.
824. L. A. LYUSTERNIK (1934), On conditional extrema of functionals, *Math. Sbornik* **41**, 390–401.
825. U. MACKENROTH (1982), Convex parabolic boundary control problems with pointwise state constraints, *J. Math. Anal. Appl.* **87**, 256–277.
826. G. G. MAGARIL-IL'YAEV (1978), An implicit function theorem for Lipschitzian mappings, *Russian Math. Surveys* **33**, 209–210.
827. E. N. MAHMUDOV (2005), On duality in problems of optimal control described by convex differential inclusions of Goursat-Darboux type, *J. Math. Anal. Appl.* **307**, 628–640.
828. E. N. MAHMUDOV (2005), The optimality principle for discrete and first order partial differential inclusions, *J. Math. Anal. Appl.* **308**, 605–619.
829. V. L. MAKAROV, M. J. LEVIN AND A. M. RUBINOV (1995), *Mathematical Economic Theory: Pure and Mixed Types of Economic Mechanisms*, North-Holland, Amsterdam, The Netherlands.
830. K. MALANOWSKI (1969), On optimal control of the vibrating string, *SIAM J. Control* **7**, 260–271.
831. K. MALANOWSKI (1979), On convergence of finite difference approximations to control and state constrained convex optimal control problems, *Arch. Autom. Telemech.* **24**, 319–337.
832. K. MALANOWSKI (1987), *Stability of Solutions to Convex Problems of Optimization*, Springer, Berlin.
833. K. MALANOWSKI (1994), Regularity of solutions in stability analysis of optimization and optimal control problems, *Control Cybernet.* **23**, 61–86.
834. K. MALANOWSKI (2003), Two-norm approach in stability and sensitivity analysis of optimization and optimal control problems, *Advan. Math. Sci. Appl.* **2**, 397–443.

835. K. MALANOWSKI AND Y. SOKOLOWSKI (1986), Sensitivity of solutions to convex, control constrained optimal control problems for distributed parameter systems, *J. Math. Anal. Appl.* **120**, 240–263.

836. G. G. MALCOLM AND B. S. MORDUKHOVICH (2001), Pareto optimality in nonconvex economies with infinite-dimensional commodity spaces, *J. Global Optim.* **20**, 323–346.

837. M. MALISOFF, L. RIFFORD AND E. SONTAG (2003), Remarks on input to state stabilization, *Proc. 42rd IEEE Conf. Dec. Cont.*, pp. 1053–1058, Maui, Hawaii.

838. M. MALISOFF, L. RIFFORD AND E. SONTAG (2004), Global asymptotic controllability implies input to state stabilization, *SIAM J. Control Optim.* **42**, 1121–1138.

839. O. L. MANGASARIAN (1985), A condition number for differentiable convex inequalities, *Math. Oper. Res.* **10**, 175–179.

840. O. L. MANGASARIAN (1994), *Nonlinear Programming*, SIAM Publications, Philadelphia, Pennsylvania.

841. O. L. MANGASARIAN AND S. FROMOVITZ (1967), The Fritz John condition in the presence of equality and inequality constraints, *J. Math. Anal. Appl.* **17**, 34–47.

842. O. L. MANGASARIAN AND T. H. SHIAU (1986), Error bounds for monotone linear complementarity problems, *Math. Progr.* **36**, 81–89.

843. K. B. MANSIMOV (1998), Singular controls in systems of neutral type, *Autom. Remote Control* **59**, 653–661.

844. K. B. MANSIMOV (2001), On the theory of necessary conditions for optimality in a problem with distributed parameters, *Comput. Maths. Math. Phys.* **41**, 1429–1443.

845. C. MARCELLI (2002), Variational problems with nonconvex, noncoercive, highly discontinuous integrands: characterization and existence of minimizers, *SIAM J. Control Optim.* **40**, 1473–1490.

846. C. MARCELLI (2008), Necessary and sufficient conditions for optimality of nonconvex, noncoercive autonomous variational problems with constraints, *Trans. Amer. Math. Soc.* **360**, 5201–5227.

847. C. MARCELLI, E. OUTKINE AND M. SYTCHEV (2002), Remarks on necessary conditions for minimizers of one-dimensional variational problems, *Nonlinear Anal.* **48**, 979–993.

848. S. MARCELLIN (2004), *Intégration D'epsilon-sous-différentiels et Problémes D'évolution Non Convexes*, Ph.D. dissertation, Department of Mathematics, University of Montpellier II, Montpellier, France.

849. S. MARCELLIN AND L. THIBAULT (2005), Integration of ε-Fenchel subdifferentials and maximal cyclic monotonicity, *J. Global Optim.* **32**, 83–91.

850. A. MARCHAUD (1934), Sur les champs de demi-cônes et les équations différentielles du premier ordre, *Bull. Sc. Math.* **62**, 12–38.

851. A. MARINO AND M. TOSQUES (1990), Some variational problems with lack of convexity and some partial differential inequalities, in *Methods of Nonconvex Analysis*, edited by A. Cellina, Lecture Notes in Mathematics **1446**, pp. 58–83, Springer, Berlin.

852. J. E. MARTINEZ-LEGAZ AND P. H. SACH (1999), A new subdifferential in quasiconvex analysis, *J. Convex Anal.* **6**, 1–12.

853. J. E. MARTINEZ-LEGAZ AND M. THÉRA (1996), Epsilon-subdifferentials in terms of subdifferentials, *Set-Valued Anal.* **4**, 327–332.

854. A. Mas-Colell (1985), Pareto optima and equilibria: The inifinite-dimensional case, in *Advances in Economics*, edited by C. D. Aliprantis et al., pp. 25–42, Springer, New York.
855. A. Mas-Colell (1986), Valuation equilibrium and Pareto optimum revisited, in *Contributions to Mathematical Economics*, edited by W. Hildenbrand and A. Mas-Collel, pp. 317–331, North-Holland, Amsterdam, The Netherlands.
856. A. Mas-Colell, M. D. Whinston and J. R. Green (1995), *Microeconomic Theory*, Oxford University Press, Oxford, UK.
857. H. Maurer and N. P. Osmolovskii (2003), Second order conditions, *Control Cybernet.* **32**, 555–584.
858. H. Maurer and J. Zowe (1979), First and second-order necessary and sufficient optimality conditions for infinite-dimensional programming problems, *Math. Progr.* **16**, 98–110.
859. A. Mayer (1884), Zur aufstellung der kriterien des maximums und minimums der einfachen integrale bei variabeln grenwerten, *Leipziger Berichte* **36**, 99–128.
860. E. J. McShane (1939), On multipliers for Lagrange problem, *Amer. J. Math.* **91**, 809–819.
861. E. J. McShane (1940), Generalized curves, *Duke Math. J.* **6**, 513–536.
862. E. J. McShane (1940), Necessary conditions for generalized curve problems in the calculus of variations, *Duke Math. J.* **7**, 1–27.
863. E. J. McShane (1967), Relaxed controls and vbariational problems, *SIAM J. Contr.* **5**, 438–485.
864. E. J. McShane (1973), The Lagrange multiplier rule, *Amer. Math. Monthly* **80**, 922–925.
865. E. J. McShane (1989), The calculus of variations from the beginning through optimal control theory, *SIAM J. Control Optim.* **27**, 916–939.
866. N. G. Medhin (1990), Minimizing sequences for differential inclusion problems in the presence of state constraints, *Comput. Math. Appl.* **19**, 127–134.
867. N. G. Medhin (1995), On optimal control of functional-differential systems, *J. Optim. Theory Appl.* **85**, 363–376.
868. T. K. Melikov (1996), An analogue of Pontryagin's maximum principle in systems with aftereffect of neutral type, *Comput. Maths. Math. Phys.* **26**, 1541–1546.
869. A. A. Melikyan (1998), *Generalized Characteristics of First Order PDEs*, Birkhäuser, Boston, Massachusetts.
870. P. Michel and J.-P. Penot (1984), Calcul sous-différentiel pour des fonctions lipschitziennes et non-lipschitziennes, *C. R. Acad. Sci. Paris* **298**, 684–687.
871. P. Michel and J.-P. Penot (1992), A generalized derivative for calm and stable functions, *Diff. Integ. Eq.* **5**, 189–196.
872. R. Mifflin (1977), Semismooth and semiconvex functions in constrained optimization, *SIAM J. Control Optim.* **15**, 957–972.
873. R. Mifflin and C. Sagastizábal (2003), Primal-dual gradient structured functions: Second-order results; links to epi-derivatives and partly smooth functions, *SIAM J. Optim.* **13**, 1174–1194.
874. R. Mifflin and C. Sagastizábal (2004), $\mathcal{VU}$-smoothness and proximal point results for some nonconvex functions, *Optim. Meth. Soft.* **19**, 463-478.

875. A. A. MILYUTIN (1999), Convex-valued Lipschitz differential inclusions and Pontryagin's maximum principle, in *Optimal Control*, edited by R. V. Gamkrelidge, Vol. 4, pp. 175–187, VINITI, Moscow.
876. A. A. MILYUTIN (2003), On strengthening the conditions by Clarke and Smirnov for convex-valued differential inclusions, *Math. Sbornik* **194**, 261–280.
877. A. A. MILYUTIN AND N. P. OSMOLOVSKII (1998), *Calculus of Variations and Optimal Control*, American Mathematical Society, Providence, Rhode Island.
878. L. I. MINCHENKO (1999), Necessary optimality conditions for differential-difference inclusions, *Nonlinear Anal.* **35**, 307–322.
879. L. I. MINCHENKO (2003), Multivalued analysis and differential properties of multivalued mappings and marginal functions, *J. Math. Sci.* **116**, 3266–3302.
880. L. I. MINCHENKO AND S. SIROTKO (2002), Controllability of nonsmooth discrete systems with delays, *Optimization* **51**, 161–174.
881. L. I. MINCHENKO AND A. A. VOLOSEVICH (2003), Value function and necessary optimality conditions in optimal control problems for differential-difference inclusions, *Nonlinear Anal.* **53**, 407–424.
882. H. MINKOWSKI (1911), *Theorie der Konvexen Körper, Insbesondere Begründung inhres Ober Flächenbegriffs*, Gesammelte Abhandlungen, II, B. G. Teubner, Leipzig.
883. V. J. MIZEL AND T. I. SEIDMAN (1997), An abstract bang-bang principle and time-optimal boundary control of the heat equation, *SIAM J. Control Optim.* **35**, 1204–1216.
884. N. N. MOISEEV (1971), *Numerical Methods in Optimal Systems Theory*, Nauka, Moscow.
885. N. N. MOISEEV (1975), *Elements of Optimal Systems Theory*, Nauka, Moscow.
886. M. D. P. MONTEIRO MARQUES (1993), *Differential Inclusions in Nonsmooth Mechanical Problems*, Birkhäuser, Boston, Massachusetts.
887. B. S. MORDUKHOVICH (1976), Maximum principle in problems of time optimal control with nonsmooth constraints, *J. Appl. Math. Mech.* **40**, 960–969.
888. B. S. MORDUKHOVICH (1977), Existence of optimal controls, *J. Soviet Math.* **7**, 850–886.
889. B. S. MORDUKHOVICH (1977), Approximation and maximum principle for nonsmooth problems of optimal control, *Russian Math. Surveys* **196**, 263–264.
890. B. S. MORDUKHOVICH (1978), On difference approximations of optimal control systems, *J. Appl. Math. Mech.* **42**, 452–461.
891. B. S. MORDUKHOVICH (1978), On the theory of difference approximations of continuous-time control systems, in *Third USSR Conference on Operations Research*, edited by N. N. Moiseev, pp. 330–331, Gorky, USSR.
892. B. S. MORDUKHOVICH (1980), Metric approximations and necessary optimality conditions for general classes of extremal problems, *Soviet Math. Dokl.* **22**, 526–530.
893. B. S. MORDUKHOVICH (1981), Penalty functions and necessary conditions for an extremum in nonsmooth and nonconvex optimization problems, *Russian Math. Surveys* **36**, 242–243.
894. B. S. MORDUKHOVICH (1984), Nonsmooth analysis with nonconvex generalized differentials and adjoint mappings, *Dokl. Akad. Nauk BSSR* **28**, 976–979.

895. B. S. MORDUKHOVICH (1984), Duality theory in systems with aftereffect, *J. Appl. Math. Mech.* **48**, 440–447.
896. B. S. MORDUKHOVICH (1984), Controllability, observability and duality in dynamical systems with aftereffect. *Autom. Remote Control* **45**, 1019–1027.
897. B. S. MORDUKHOVICH (1985), On necessary conditions for an extremum in nonsmooth optimization, *Soviet Math. Dokl.* **32**, 215–220.
898. B. S. MORDUKHOVICH (1986), Optimal control of the groundwater regime on engineering reclamation systems, *Water Resources* **12**, 244–253.
899. B. S. MORDUKHOVICH (1986), On the theory of difference approximations in optimal control, *Dokl. Akad. Nauk BSSR* **30**, 1964–1067.
900. B. S. MORDUKHOVICH (1987), Approximation methods and optimality conditions in nonconvex control problems, *Soviet Math. Dokl.* **36**, 164–168.
901. B. S. MORDUKHOVICH (1988), *Approximation Methods in Problems of Optimization and Control*, Nauka, Moscow.
902. B. S. MORDUKHOVICH (1988), Approximation and optimization of differential inclusions, *Cybernetics* **24**, 781–788.
903. B. S. MORDUKHOVICH (1988), Approximate maximum principle for finite-difference control systems, *USSR Comput. Maths. Math. Phys.* **28**, 106–114.
904. B. S. MORDUKHOVICH (1989), Necessary optimality conditions for nonsmooth control systems with free time, *Diff. Eq.* **25**, 290–299.
905. B. S. MORDUKHOVICH (1990), Minimax design for a class of distributed parameter systems, *Autom. Remote Control* **50**, 262–283.
906. B. S. MORDUKHOVICH (1990), Maximum principle for nonconvex finite difference systems, in *Analysis and Optimization of Systems*, edited by A. Bensoussan and J.-L. Lions, Lecture Notes Cont. Inf. Sci. **144**, pp. 539–548, Springer, Berlin.
907. B. S. MORDUKHOVICH (1992), Sensitivity analysis in nonsmooth optimization, in *Theoretical Aspects of Industrial Design*, edited by D. A. Field and V. Komkov, SIAM Proc. Appl. Math. **58**, pp. 32–46, Philadelphia, Pennsylvania.
908. B. S. MORDUKHOVICH (1992), On variational analysis of differential inclusions, in *Optimization and Nonlinear Analysis*, edited by A. Ioffe, L. Marcus and S. Reich, Pitman Researh Notes Math. Ser. **244**, pp. 199–213, Longman, Harlow, Essex, UK.
909. B. S. MORDUKHOVICH (1993), Complete characterization of openness, metric regularity, and Lipschitzian properties of multifunctions, *Trans. Amer. Math. Soc.* **340**, 1–35.
910. B. S. MORDUKHOVICH (1994), Generalized differential calculus for nonsmooth and set-valued mappings, *J. Math. Anal. Appl.* **183**, 250–288.
911. B. S. MORDUKHOVICH (1994), Lipschitzian stability of constraint systems and generalized equations, *Nonlinear Anal.* **22**, 173–206.
912. B. S. MORDUKHOVICH (1994), Stability theory for parametetic generalized equations and variational inequalities via nonsmooth analysis, *Trans. Amer. Math. Soc.* **343**, 609–658.
913. B. S. MORDUKHOVICH (1994), Sensitivity analysis for constraint and variational systems by means of set-valued differentiation, *Optimization* **31**, 13–46.
914. B. S. MORDUKHOVICH (1994), Necesary optimality and controllability conditions for nonsmooth control systems, in *Proc. 33rd IEEE Conf. Dec. Cont.*, pp. 3992–3997, Orlando, Florida.

915. B. S. MORDUKHOVICH (1995), Discrete approximations and refined Euler-Lagrange conditions for nonconvex differential inclusions, *SIAM J. Control Optim.* **33**, 882–915.

916. B. S. MORDUKHOVICH (1996), Optimization and finite difference approximations of nonconvex differential inclusions with free time, in *Nonsmooth Analysis and Geometric Methods in Deterministic Optimal Control*, edited by B. S. Mordukhovich and H. J. Sussmann, pp. 153–202, Springer, New York.

917. B. S. MORDUKHOVICH (1997), Coderivatives of set-valued mappings: Calculus and applications, *Nonlinear Anal.* **30**, 3059–3070.

918. B. S. MORDUKHOVICH (1999), Minimax design of constrained parabolic systems, in *Control of Distributed Parameter and Stochastic Systems*, edited by S. Chen et al., pp. 111–118, Kluwer, Boston, Massachusetts.

919. B. S. MORDUKHOVICH (1999), On variational analysis in infinite dimensions, in *Systems Modelling and Optimization*, edited by M. P. Polis et al., pp. 189–197, Chapman and Hall/CRC Press, Boca Raton, Florida.

920. B. S. MORDUKHOVICH (2000), Abstract extremal principle with applications to welfare economics, *J. Math. Anal. Appl.* **251**, 187–216.

921. B. S. MORDUKHOVICH (2000), Optimal control of difference, differential, and differential-difference inclusions, *J. Math. Sci.* **100**, 2613–2632.

922. B. S. MORDUKHOVICH (2001), The extremal principle and its applications to optimization and economics, in *Optimization and Related Topics*, edited by A. Rubinov and B. Glover, Applied Optimization **47**, pp. 343–369, Kluwer, Dordrecht, The Netherlands.

923. B. S. MORDUKHOVICH (2002), Calculus of second-order subdifferentials in infinite dimensions, *Control Cybernet.* **31**, 557–573.

924. B. S. MORDUKHOVICH (2004), Coderivative analysis of variational systems, *J. Global Optim.* **28**, 347–362.

925. B. S. MORDUKHOVICH (2004), Necessary conditions in nonsmooth minimization via lower and upper subgradients, *Set-Valued Anal.* **12**, 163–193.

926. B. S. MORDUKHOVICH (2004), Equilibrium problems with equilibrium constraints via multiobjective optimization, *Optim. Meth. Soft.* **19**, 479–492.

927. B. S. MORDUKHOVICH (2004), Lipschitzian stability of parametric constraint systems in infinite dimensions, in *Generalized Convexity, Generalized Monotonicity and Aplications*, edited by A. Eberhard et al., Nonconvex Optimization and Its Applications, pp. 39–59, Springer, New York.

928. B. S. MORDUKHOVICH (2005), Optimization and equilibrium problems with equilibrium constraints, *OMEGA* **33**, 379–384.

929. B. S. MORDUKHOVICH (2005), Sensitivity analysis for variational systems, in *Variational Analysis and Applications*, edited by F. Giannessi and A. Maugeri, pp. 723–743, Springer, Berlin.

930. B. S. MORDUKHOVICH (2005), Nonlinear prices in nonconvex economies with classical Pareto and strong Pareto optimal allocations, *Positivity* **9**,541–568.

931. B. S. MORDUKHOVICH (2005), Sensitivity analysis for generalized variational and hemivariational inequalities, in *Advances in Analysis*, edited by H. G. W. Begehr et al., pp. 305–314, World Scientific Publishing, London, UK.

932. B. S. MORDUKHOVICH (2005), Optimal control of evolution inclusions, *Nonlinear Anal.*, **63**, 775–784.

933. B. S. MORDUKHOVICH (2006), Coderivative calculus and robust Lipschitzian stability of variational systems, *J. Convex Anal.*, **13**, 799–822.

934. B. S. MORDUKHOVICH AND N. M. NAM (2005), Variational stability and marginal functions via generalized differentiation, *Math. Oper. Res.* **30**, 1–18.
935. B. S. MORDUKHOVICH AND N. M. NAM (2005), Subgradients of distance functions with some applications, *Math. Progr.*, **104**, 635–668.
936. B. S. MORDUKHOVICH AND N. M. NAM (2006), Subgradients of distance functions at out-of-state points, *Taiwan. J. Math.*, **10**, 299–326.
937. B. S. MORDUKHOVICH, N. M. NAM AND N. D. YEN (2009), Subgradients of marginal functions in parametric mathematical programming, *Math. Progr.*, **116**, 369–396.
938. B. S. MORDUKHOVICH, N. M. NAM AND N. D. YEN (2006), Fréchet subdifferential calculus and optimality conditions in mathematical programming, *Optimization*, **55**, 685–708.
939. B. S. MORDUKHOVICH AND J. V. OUTRATA (2001), Second-order subdifferentials and their applications, *SIAM J. Optim.* **12**, 139–169.
940. B. S. MORDUKHOVICH, J. V. OUTRATA AND M. ČERVINKA (2007), Equilibrium problems with complementarity constraints: Case study with applications to oligopolistic markets, *Optimization*, **56**, 479–494.
941. B. S. MORDUKHOVICH AND T. PENNANEN (2007), Epi-convergent approximations for generalized Bolza problems, *Optim. Lett.* **1**, 379–390.
942. B. S. MORDUKHOVICH AND V. M. RAKETSKII (1980), Necessary optimality conditions of the maximum principle type in discrete approximations of nonconvex control problems with constraints on trajectories, *Vest. Beloruss. Univ.*, Ser. 1, No. 3, 72–73; Depon. VINITI #791-79, Moscow.
943. B. S. MORDUKHOVICH AND J.-P. RAYMOND (2004), Dirichlet boundary control of hyperbolic equations in the presence of state constraints, *Appl. Math. Optim.* **49**, 145–157.
944. B. S. MORDUKHOVICH AND J.-P. RAYMOND (2004), Neumann boundary control of hyperbolic equations with pointwise state constraints, *SIAM J. Control Optim.* **43**, 1354–1372.
945. B. S. MORDUKHOVICH AND A. M. SASONKIN (1985), Duality and optimality conditions in control problems for functional-differential systems of neutral type, *Diff. Eq.* **21**, 532–340.
946. B. S. MORDUKHOVICH AND Y. SHAO (1995), Differential characterizations of covering, metric regularity, and Lipschitzian properties of multifunctions between Banach spaces, *Nonlinear Anal.* **25**, 1401–1424.
947. B. S. MORDUKHOVICH AND Y. SHAO (1995), On nonconvex subdifferential calculus in Banach spaces, *J. Convex Anal.* **2**, 211–227.
948. B. S. MORDUKHOVICH AND Y. SHAO (1996), Extremal characterizations of Asplund spaces, *Proc. Amer. Math. Soc.* **124**, 197–205.
949. B. S. MORDUKHOVICH AND Y. SHAO (1996), Nonsmooth sequential analysis in Asplund spaces, *Trans. Amer. Math. Soc.* **348**, 1235–1280.
950. B. S. MORDUKHOVICH AND Y. SHAO (1996), Nonconvex coderivative calculus for infinite-dimensional multifunctions, *Set-Valued Anal.* **4**, 205–236.
951. B. S. MORDUKHOVICH AND Y. SHAO (1997), Stability of multifunctions in infinite dimensions: Point criteria and applications, *SIAM J. Control Optim.* **35**, 285–314.
952. B. S. MORDUKHOVICH AND Y. SHAO (1997), Fuzzy calculus for coderivatives of multifunctions, *Nonlinear Anal.* **29**, 605–626.
953. B. S. MORDUKHOVICH AND Y. SHAO (1998), Mixed coderivatives of set-valued mappings in variational analysis, *J. Appl. Anal.* **4**, 269–294.

954. B. S. MORDUKHOVICH, Y. SHAO AND Q. J. ZHU (2000), Viscosity coderivatives and their limiting behavior in smooth Banach spaces, *Positivity* **4**, 1–39.
955. B. S. MORDUKHOVICH AND I. A. SHVARTSMAN (2002), Discrete maximum principle for nonsmooth optimal control problems with delays, *Cybernet. Systems Anal.* **38**, 255–264.
956. B. S. MORDUKHOVICH AND I. A. SHVARTSMAN (2004), The approximate maximum principle in constrained optimal control, *SIAM J. Control Optim.* **43**, 1037–1062.
957. B. S. MORDUKHOVICH AND I. A. SHVARTSMAN (2004), Optimization and feedback control of constrained parabolic systems under uncertain perturbations, in *Optimal Control, Stabilization and Nonsmooth Analysis*, edited by M. de Queiroz, M. Malisoff and P. Wolenski, Lecture Notes Cont. Inf. Sci. **301**, pp. 121–132, Springer, New York.
958. B. S. MORDUKHOVICH, J. S. TREIMAN AND Q. J. ZHU (2003), An extended extremal principle with applications to multiobjective optimization, *SIAM J. Optim.* **14**, 359–379.
959. B. S. MORDUKHOVICH AND R. TRUBNIK (2001), Stability of discrete approximations and necessary optimality conditions for delay-differential inclusions, *Ann. Oper. Res.* **101**, 149–170.
960. B. S. MORDUKHOVICH AND B. WANG (2000), On variational characterizations of Asplund spaces, in *Constructive, Experimental and Nonlinear Analysis*, edited by M. Théra, Canad. Math. Soc. Conf. Proc. **27**, 245–254.
961. B. S. MORDUKHOVICH AND B. WANG (2001), Sequential normal compactness in variational analysis, *Nonlinear Anal.* **47**, 717–728.
962. B. S. MORDUKHOVICH AND B. WANG (2002), Necessary optimality and suboptimality conditions in nondifferentiable programming via variational principles, *SIAM J. Control Optim.* **41**, 623–640.
963. B. S. MORDUKHOVICH AND B. WANG (2002), Extensions of generalized differential calculus in Asplund spaces, *J. Math. Anal. Appl.* **272**, 164–186.
964. B. S. MORDUKHOVICH AND B. WANG (2003), Calculus of sequential normal compactness in variational analysis, *J. Math. Anal. Appl.* **282**, 63–84.
965. B. S. MORDUKHOVICH AND B. WANG (2003), Differentiability and regularity of Lipschitzian mappings, *Proc. Amer. Math. Soc.* **131**, 389–399.
966. B. S. MORDUKHOVICH AND B. WANG (2004), Generalized differentiation for moving objects, in *Optimal Control, Stabilization and Nonsmooth Analysis*, edited by M. de Queiroz, M. Malisoff and P. Wolenski, Lecture Notes Cont. Inf. Sci. **301**, pp. 351-361, Springer, New York.
967. B. S. MORDUKHOVICH AND B. WANG (2004), Restrictive metric regularity and generalized differential calculus in Banach spaces, *Int. J. Maths. Math. Sci.* **50**, 2650–2683.
968. B. S. MORDUKHOVICH AND B. WANG (2005), Restrictive metric regularity in variational analysis, *Nonlinear Anal.*, **63**, 805–811.
969. B. S. MORDUKHOVICH AND B. WANG (2008), Generalized differential and normal compactness calculi for moving objects, *Optimization* **57**, 17–40.
970. B. S. MORDUKHOVICH AND D. WANG (2005), Optimal control of semilinear unbounded differential inclusions, *Nonlinear Anal.*, **63**, 847–853.
971. B. S. MORDUKHOVICH AND D. WANG (2005), Optimal control of semilinear evolution inclusions via discrete approximations, *Control Cybernet.* **34**, 849–870.

972. B. S. Mordukhovich and L. Wang (2002), Optimal control of hereditary differential inclusions, in *Proc. 41st IEEE Conf. Decis. Control*, pp. 1107–1112, Las Vegas, Nevada.
973. B. S. Mordukhovich and L. Wang (2003), Optimal control of constrained delay-differential inclusions with multivalued initial conditions, *Control Cybernet.* **32**, 585–609.
974. B. S. Mordukhovich and L. Wang (2004), Optimal control of neutral functional-differential inclusions, *SIAM J. Control Optim.* **43**, 111–136.
975. B. S. Mordukhovich and L. Wang (2004), Optimal control of differential-algebraic inclusions, in *Optimal Control, Stabilization and Nonsmooth Analysis*, edited by M. de Queiroz, M. Malisoff and P. Wolenski, Lecture Notes Cont. Inf. Sci. **301**, pp. 73–83, Springer, New York.
976. B. S. Mordukhovich and L. Wang (2005), Optimal control of delay systems with differential and algebraic dynamic constraints, *ESAIM: Control Optim. Calc. Var.* **11**, 285–309.
977. B. S. Mordukhovich and L. Wang (2005), Optimal control of nonautonomous functional-differential inclusions of neutral type, *Nonlinear Anal.*, to appear.
978. B. S. Mordukhovich and K. Zhang (1997), Minimax control of parabolic equations with Dirichlet boundary conditions and state constraints, *Appl. Math. Optim.* **36**, 323–360.
979. B. S. Mordukhovich and K. Zhang (1998), Dirichlet boundary control of parabolic systems with pointwise state constraints, *Int. Ser. Numer. Math.* **126**, 223–236.
980. J.-J. Moreau (1962), Fonctions convexes duales et points proximaux dans un espace hilbertien, *C. R. Acad. Sci. Paris* **255**, 2897–2899.
981. J.-J. Moreau (1963), Fonctionelles sous-différentiables, *C. R. Acad. Sci. Paris* **257**, 4117–4119.
982. J.-J. Moreau (1965), Proximité dualité dans espace hilbertien, *Bull. Soc. Math. France* **93**, 273–299.
983. J.-J. Moreau (1977), Evolution problems associated with moving coinvex set in a Hilbert space, *J. Diff. Eq.* **26**, 347–374.
984. J.-J. Moreau (1988), Bounded variation in time, in *Topics in Nonsmooth Mechanics*, edited by J.-J. Moreau, P. D. Panagiotopoulos and G. Strang, pp. 1–74, Birkhäuser, Boston, Massachusetts.
985. D. Motreanu and N. Pavel (1999), *Tangency, Flow-Invariance for Differential Equations, and Optimization Problems*, Marcel Dekker, New York.
986. M. Moussaoui and A. Seeger (1994), Sensitivity analysis of optimal value functions of convex parametric programs with possibly empty solution sets, *SIAM J. Optim.* **3**, 659–675.
987. M. Moussaoui and A. Seeger (1996), Epsilon-maximum principle of Pontryagin type and perturbation analysis of convex optimal control problems, *SIAM J. Control Optim.* **34**, 407–427.
988. M. Mureşan (2001), *Analiză Nenetedăsi Aplicatii*, Risoprint, Cluj-Napoca, Romania.
989. A. D. Myshkis (1951), *Linear Differential Equations with Retarded Argument*, GITTL, Moscow.
990. N. Nagumo (1942), Ber diw lage der integralkurven gewohlicher differenialgleichungen, *Proc. Phys. Math. Soc. Japan* **24**, 551–559.

991. N. M. NAM AND N. D. YEN (2007), Relationships between approximate Jacobians and coderivatives, *J. Nonlinear Convex Anal.*, **8**,121-133

992. I. NAMIOKA AND R. R. PHELPS (1975), Banach spaces which are Asplund spaces, *Duke Math. J.* **42**, 735–750.

993. Z. NANIEWICZ (2007), Pseudomonotonicity and economic equilibrium problem in reflexive Banach space, *Math. Oper. Res.* **32**, 436-466.

994. Z. NANIEWICZ AND P. D. PANAGIOTOPOULOS (1995), *Mathematical Theory of Hemivariational Inequalities and Applications*, Marcel Dekker, New York.

995. J. F. NASH (1951), Non-cooperative games, *Annals Math.* **54**, 286–295.

996. M. H. NAYAKKANKUPPAM AND M. L. OVERTON (1999), Conditioning of semilinear programs, *Math. Progr.* **85**, 525–540.

997. P. NEITTAAANMÄKI AND D. TIBA (1994), *Optimal Control of Nonlinear Parabolic Systems*, Marcel Dekker, New York.

998. Y. E. NESTEROV (2005), Lexicographical differentiation of nonsmooth functions, *Math. Progr.*, **104**, 669–700.

999. Y. E. NESTEROV AND A. S. NEMIROVSKY (1993), *Interior Point Methods in Convex Optimization: Theory and Applications*, SIAM Publications, Philadelphia, Pennsylvania.

1000. J. VON NEUMANN AND O. MORGENSTERN (1944), *Theory of Games and Economic Behaviour*, Princeton University Press, Princeton, New Jersey.

1001. L. W. NEUSTADT (1966), An abstract variational theory with applications to a broad class of optimization problems, I: General theory, *SIAM J. Control* **4**, 505–527.

1002. L. W. NEUSTADT (1976), *Optimization: a Theory of Necessary Conditions*, Princeton University Press, Princeton, New Jersey.

1003. K. F. NG AND W. H. YANG (2004), Regularities and their relations to error bounds, *Math. Progr.* **99**, 521–538.

1004. K. F. NG AND X. Y. ZHENG (2001), Error bounds for lower semicontinuous functions in normed spaces, *SIAM J. Optim.* **12**, 1–17.

1005. K. F. NG AND X. Y. ZHENG (2003), Global weak sharp minima in Banach spaces, *SIAM J. Control Optim.* **41**, 1858–1885.

1006. H. V. NGAI, D. T. LUC AND M. THÉRA(2000), Approximate convex functions, *J. Nonlinear Convex Anal.* **1**, 155–176.

1007. H. V. NGAI, D. T. LUC AND M. THÉRA (2002), Extensions of Fréchet ε-subdifferential calculus and applications, *J. Math. Anal. Appl.* **268**, 266–290.

1008. N. V. NGAI AND M. THÉRA (2001), Metric regularity, subdifferential calculus and applications, *Set-Valued Anal.* **9**, 187–216.

1009. N. V. NGAI AND M. THÉRA (2002), A fuzzy necessary optimality condition for non-Lipschitz optimization in Asplund spaces, *SIAM J. Optim.* **12**, 656–668.

1010. N. V. NGAI AND M. THÉRA (2004), Error bounds, and implicit multifunction theorem in smooth Banach spaces and applications to optimization, *Set-Valued Anal.* **12**, 195–223.

1011. A. NIJENHUIS (1974), Strong derivatives and inverse mappings, *Amer. Math. Monthly* **81**, 969–980.

1012. M. S. NIKOLSKII (2004), Difference approximation of some optimization problems, *Comput. Maths. Math. Phys.* **44**, 467–475.

1013. K. NITKA-STYCZEN (1988), Difference approximations of optimal periodic control problems, *Int. J. Control* **47**, 1893–1904.

1014. K. NITKA-STYCZEN (1989), Approximate discrete maximum principle for the discrete approximation of optimal periodic control problems, *Int. J. Control* **50**, 1863–1871.

1015. K. NITKA-STYCZEN (1999), *Optimal Periodic Control of Hereditary Processes. Theory, Algorithms and Applications*, Technical University of Wrocław, Wrocław, Poland.

1016. A. NOWAKOWSKI AND I. NOWAKOWSKA (2008), Dirichlet problem for semilinear hyperbolic equation. Multidimensional case, *J. Math. Anal. Appl.* **338**, 771-783.

1017. E. A. NURMINSKII (1979), *Methods of Solutions to Deterministic and Stochastic Minimax Problems*, Naukova Dumka, Kiev.

1018. M. N. OĞUZTÖRELI (1966), *Time-Lag Control Systems*, Academic Press, New York.

1019. A. OLBROT (1976), Control of retarded systems with function space constraints: Necessary optimality conditions, *Control Cybernet.* **5**, 5–31.

1020. C. OLECH (1976), Existence theory in optimal control problems, in *Control Theory and Topics in Functional Analysis*, Vol. 1, pp. 291–328, International Atomic Energy Agency, Vienna.

1021. N. ORTIZ (2005), Necessary conditions for the neutral problem of Bolza with continuously varying time delay, *J. Math. Anal. Appl.* **305**, 513–527.

1022. N. ORTIZ AND P. R. WOLENSKI (2004), The decoupling technique for continuously varying time delay systems, *Set-Valued Anal.* **12**, 225–239.

1023. Y. S. OSIPOV, L. PANDOLFI AND V. I. MAKSIMOV (2001), Problems of dynamical reconstruction and robust boundary control: The case of the Dirichlet boundary conditions, *J. Inverse Ill-Posed Probl.* **9**, 149–162.

1024. J. V. OUTRATA (1999), Optimality conditions for a class of mathematical programs with equilibrium constraints, *Math. Oper. Res.* **24**, 627–644.

1025. J. V. OUTRATA (1999), Optimality conditions for a class of mathematical programs with equilibrium constraints: Strongly regular case, *Kybernetika* **35**, 177–193.

1026. J. V. OUTRATA (2000), On mathematical programs with complementarity constraints, *Optim. Meth. Soft.* **14**, 117–137.

1027. J. V. OUTRATA (2000), A generalized mathematical program with equilibrium constraints, *SIAM J. Control Optim.* **38**, 1623–1638.

1028. J. V. OUTRATA (2001), On constraint qualifications for mathematical programs with mixed complementarity constraints, in *Complementarity, Applications and Extensions*, edited by M. C. Ferris, O. L. Mangasarian and J.-S. Pang, pp. 253–272, Kluwer, Dordrecht, The Netherlands.

1029. J. V. OUTRATA (2004), A note on a class of equilibrium problems with equilibrium constraints, *Kybernetika* **40**, 585–594.

1030. J. V. OUTRATA (2005), Mathematical programs with equilibrium constraints: Theory and numerical methods, in *Nonsmooth Mechanics of Solids, CISM Lecture Notes*, edited by J. Haslinger and G. E. Stavroulakis, Springer, New York.

1031. J. V. OUTRATA, M. KOČVARA AND J. ZOWE (1998), *Nonsmooth Approach to Optimization Problems with Equilibrium Constraints*, Kluwer, Dordrecht, The Netherlands.

1032. J. V. OUTRATA AND W. RÖMISCH (2005), On optimality conditions for some nonsmooth optimization problems over L^p spaces, *Optim. Theory Appl.* **126**, 1–28.

1033. M. L. Overton (1988), On minimizing the maximum eigenvalue of a symmetric matrix, *SIAM J. Matrix Anal. Appl.* **9**, 256–268.
1034. M. L. Overton and R. S. Womersley (1993), Optimality conditions and duality theory for minimizing sums of the largest eigenvalue of symmetric matrices, *Math. Progr.* **62**, 321–337.
1035. A. E. Ozdaglar and D. P. Bertsekas (2004), The relation between pseudonormality and quasiregularity in constrained optimization, *Optim. Meth. Soft.* **19**, 493–506.
1036. Z. Páles and V. Zeidan (1994), First and second order necessary conditions for control problems with constraints, *Trans. Amer. Math. Soc.* **346**, 421–453.
1037. Z. Páles and V. Zeidan (1996), Generalized Hessians for $C^{1,1}$ functions in infinite dimensional normed spaces, *Math. Progr.* **74**, 59–78.
1038. Z. Páles and V. Zeidan (2004), Critical and critical tangent cones in optimization problems, *Set-Valued Anal.* **12**, 241–258.
1039. D. Pallaschke, P. Recht and R. Urbański (1987), On extensions of the second-order derivative, *Bull. Polish Acad. Sci., Ser. Math.* **35**, 750–763.
1040. D. Pallaschke and S. Rolewicz (1998), *Foundations of Mathematical Optimization: Convex Analysis without Linearity*, Kluwer, Dordrecht, The Netherlands.
1041. D. Pallaschke and R. Urbański (2002), *Pairs of Compact Convex Sets*, Kluwer, Dordrecht, The Netherlands.
1042. P. D. Panagiotopoulos (1993), *Hemivariational Inequalities: Applications to Mechanics and Engineering*, Springer, Berlin.
1043. J.-S. Pang (1990), Solution differentiability and continuation of Newton's method for variational inequality problems over polyhedral sets, *J. Optim. Theory Appl.* **66**, 121–135.
1044. J.-S. Pang (1993), A degree-theoretical approach to parametric nonsmooth equations with multivalued perturbed solution sets, *Math. Progr.* **62**, 359–383.
1045. J.-S. Pang (1997), Error bounds in mathematical programming, *Math. Progr.* **79**, 299–332.
1046. J.-S. Pang and D. Stewart (2008), Differential variational inequalities, *Math. Progr.*, **113**, 345–424.
1047. J.-S. Pang and L. Qi (1993), Nonsmooth equations: Motivations and algorithms, *SIAM J. Optim.* **3**, 443–465.
1048. C. Pantelides, D. Gritsis, K. P. Morison and R. W. H. Sargent (1988), The mathematical modelling of transient systems using differential-algebraic equations, *Comput. Chem. Eng.* **12**, 449–454.
1049. N. S. Papageorgiou (1995), On parametric evolution inclusions of the subdifferential type with applications to optimal control problems, *Trans. Amer. Math. Soc.* **347**, 203–231.
1050. M. Papi and S. Sbaraglia (2003), Regularity properties of constrained set-valued mappings, *Nonlinear Anal.* **54**, 1337–1353.
1051. M. Papi and S. Sbaraglia (2005), Lipschitzian estimates in discrete-time constrained stochastic optimal control, *Dynamics Cont. Disc. Impuls. Syst.*, to appear.
1052. M. Papi and S. Sbaraglia (2005), Optimal asset-liability management with constraints: A dynamic programming approach, *Appl. Math. Comput.*, to appear.
1053. V. Pareto (1909), *Manuel d'Economie Politique*, Giard, Paris.

1054. G. PEANO (1892), Sur la définition de la dérivée, *Mathesis* **2**, 12–14.
1055. J. PEÑA (2000), Understanding the geometry of infeasible perturbations of a conic linear system, *SIAM J. Optim.* **10**, 534–550.
1056. J. PEÑA (2001), Conditioning of convex programs from a primal-dual perspectives, *Math. Oper. Res.* **26**, 206–220.
1057. J. PEÑA (2003), A characterization of the distance to infeasibility under structured perturbations, *Linear Algebra Appl.* **370**, 193–216.
1058. J. PEÑA (2004), Conic systems and sublinear mappings: Equivalent approaches, *Operations Research Letters* **32** (2004), 463–467.
1059. J. PEÑA (2005), On the block-structured distance to non-surjectivity of sublinear mappings, *Math. Progr.* **103**, 561–573.
1060. T. PENNANEN (1999), Graph-convex mappings and K-convex functions, *J. Convex Anal.* **6**, 235–266.
1061. T. PENNANEN (2005), Epi-convergent discretizations of multistage stochastic programs, *Math. Oper. Res.* **30**, 245–256.
1062. T. PENNANEN (2009), Epi-convergent discretizations of multistage stochastic programs via integration quadratures, *Math. Progr.*, **116**, 461-479.
1063. J.-P. PENOT (1974), Sous-différentieles de fonctions numériques nonconvexes, *C. R. Acad. Paris* **278**, 1553–1555.
1064. J.-P. PENOT (1978), Calcul sous-différentiel et optimisation, *J. Funct. Anal.* **27**, 248–276.
1065. J.-P. PENOT (1981), A characterization of tangential regularity, *Nonlinear Anal.* **5**, 625–643.
1066. J.-P. PENOT (1989), Metric regularity, openness and Lipschitzian behavior of multifunctions, *Nonlinear Anal.* **13**, 629–643.
1067. J.-P. PENOT (1994), Sub-hessians, super-hessions and conjugation, *Nonlinear Anal.* **23**, 689–702.
1068. J.-P. PENOT (1995), Inverse function theorems for mappings and multimappings, *Southeast. Asian Bull. Math.* **19**, 1–16.
1069. J.-P. PENOT (1996), Favorable classes of mappings and multimappings in nonlinear analysis and optimization, *J. Convex Anal.* **3**, 97–116.
1070. J.-P. PENOT (1997), Metric estimates for the calculus of multimappings, *Set-Valued Anal.* **5**, 291–308.
1071. J.-P. PENOT (1998), Compactness properties, openness criteria and coderivatives, *Set-Valued Anal.* **6**, 363–380.
1072. J.-P. PENOT (1998), Are generalized derivatives useful for generalized convex functions?, *Generalized Convexity, Generalized Monotonicity: Recent Results*, edited by J.-P. Crouzeix, J. E. Martinez-Legaz and M. Volle, pp. 3–59, Kluwer, Dordrecht, The Netherlands.
1073. R. R. PHELPS (1993), *Convex Functions, Monotone Operators and Differentiability*, 2nd edition, Springer, Berlin.
1074. M. D. R. DE PINHO (2003), Mixed constrained control problems, *J. Math. Anal. Appl.* **278**, 293–307.
1075. M. D. R. DE PINHO, M. M. A. FERREIRA AND F. A. C. C. FONTES (2002), An Euler-Lagrange inclusion for optimal control problems with state constraints, *J. Dynam. Control Systems* **8**, 23–45.
1076. M. D. R. DE PINHO, M. M. A. FERREIRA AND F. A. C. C. FONTES (2005), Unmaximized inclusion necessary conditions for nonconvex constrained optimal control problems, *ESAIM: Control Optim. Calc. Var.* **11**.

1077. M. D. R. DE PINHO AND A. ILCHMANN (2002), Weak maximum principle for optimal control problems with mixed constraints, *Nonlinear Anal.* **48**, 1179–1196.
1078. M. D. R. DE PINHO AND R. B. VINTER (1995), An Euler-Lagrange inclusion for optimal control problems, *IEEE Trans. Autom. Control* **40**, 1191–1198.
1079. M. D. R. DE PINHO AND R. B. VINTER (1997), Necessary conditions for optimal control problems involving nonlinear differential algebraic equations, *J. Math. Anal. Appl.* **212**, 493–516.
1080. M. D. R. DE PINHO, R. B. VINTER AND H. ZHENG (2001), A maximum principle for optimal control problems with mixed constraints, *IMA J. Math. Control Inf.* **18**, 189–205.
1081. S. PLASKASZ (1992), On the solution sets for differential inclusions, *Boll. Un. Mat. Ital.* **6**, 387–394.
1082. V. A. PLOTNIKOV, A. V. PLOTNIKOV AND A. N. VITYUK (1999), *Differential Equations with a Multivalued Right-Hand Side: Asymptotic Methods*, AstroPrint, Odessa, Ukraine.
1083. V. I. PLOTNIKOV (1972), Necessary and sufficient conditions for optimality and conditions for uniqueness of the optimizing functions for control systems of general form, *Izv. Akad. Nauk SSSR, Ser. Mat.* **36**, 652–679.
1084. V. I. PLOTNIKOV AND M. I. SUMIN (1982), The construction of minimizing sequences in control problems for systems with distributed paratemers, *USSR Comput. Maths. Math. Phys.* **22**, 292–296.
1085. V. I. PLOTNIKOV AND V. I. SUMIN (1972), Optimization of objects with distributed parameters that can be described by Coursat-Darboux systems, *USSR Comput. Maths. Math. Phys.* **12**, 49–57.
1086. E. POLAK (1999), *Optimization: Algorithms and Consistent Approximations*, Springer, New York.
1087. R. A. POLIQUIN (1990), Subgradient monotonicity and convex functions, *Nonlinear Anal.* **14**, 305–317.
1088. R. A. POLIQUIN (1992), An extension of Attouch's theorem and its application to second-order epi-differentiation of convexly composite functions, *Trans. Amer. Math. Soc.* **332**, 861–874.
1089. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1992), Amenable functions in optimization, in *Nonsmooth Optimization Methods and Applications*, edited by F. Giannessi, pp. 338–353, Gordon and Breach, Philadelphia, Pennsylvania.
1090. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1996), Prox-regular functions in variational analysis, *Trans. Amer. Math. Soc.* **348**, 1805–1838.
1091. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1996), Generalized Hessian properties of regularized nonsmooth functions, *SIAM J. Optim.* **6**, 1121–1137.
1092. R. A. POLIQUIN AND R. T. ROCKAFELLAR (1998), Tilt stability of a local minimum, *SIAM J. Optim.* **8**, 287–299.
1093. R. A. POLIQUIN, R. T. ROCKAFELLAR AND L. THIBAULT (2000), Local differentiability of distance functions, *Trans. Amer. Math. Soc.* **332**, 5231–5249.
1094. E. S. POLOVINKIN AND G. V. SMIRNOV (1986), An approach to differentiation of multifunctions and necessary optimality conditions for differential inclusions, *Diff. Eq.* **22**, 660–668.
1095. E. S. POLOVINKIN AND G. V. SMIRNOV (1986), Time-optimal problem for differential inclusions, *Diff. Eq.* **22**, 940–952.
1096. B. T. POLYAK (1969), Semicontinuity of intergral functionals and existence theorems on extremal problems, *Math. Sbornik* **7**, 59–77.

1097. B. T. POLYAK (1987), *Introduction to Optimization*, Optimization Software, New York.

1098. B. T. POLYAK (1998), Convexity of quadratic transformations and its use in control and optimization, *J. Optim. Theory Appl.* **99**, 553–583.

1099. B. T. POLYAK (2000), History of mathematical programming in the USSR: Analyzing the phenomenon, *Math. Progr.* **91**, 401–416.

1100. B. T. POLYAK (2001), Convexity of nonlinear image of a small ball with applicatiobs to optimization, *Set-Valued Anal.* **9**, 159–168.

1101. D. POMPEIU (1905), Fonctions de variables complexes, *Ann. Facul. Sci. Univ. Toulouse* **7**, 265–315.

1102. L. S. PONTRYAGIN, V. G. BOLTYANSKII, R. V. GAMKRELIDZE AND E. F. MISHCHENKO (1962), *The Mathematical Theory of Optimal Processes*, Wiley, New York.

1103. D. PREISS (1984), Gâteaux differentiable functions are somewhere Fréchet differentiable, *Rend. Circ. Mat. Palermo* **33**, 122–133.

1104. D. PREISS (1990), Differentiability of Lipschitz functions on Banach spaces, *J. Funct. Anal.* **91**, 312–345.

1105. A. I. PROPOI (1973), *Elements of the Theory of Optimal Discrete Processes*, Nauka, Moscow.

1106. B. N. PSHENICHNYI (1971), *Necessary Conditions for an Extremum*, Marcel Dekker, New York.

1107. B. N. PSHENICHNYI (1976), Necessary conditions for an extremum for differential inclusions, *Kibernetika* **12**, 60–73.

1108. B. N. PSHENICHNYI (1977), On necessary extremality conditions for nonsmooth functions, *Kibernetika* **13**, 92–96.

1109. B. N. PSHENICHNYI (1980), *Convex Analysis and Extremal Problems*, Nauka, Moscow.

1110. V. PTÁK (1974), A quantitative refinement of the closed graph theorem, *Czech. Math. J.* **24**, 503–506.

1111. L. QI AND J. SUN (1993), A nonsmooth version of Newton's method, *Math. Progr.* **58**, 353–368.

1112. Y. P. QIU AND T. L. MAGNANTI (1992), Sensitivity analysis for variational inequalities, *Math. Oper. Res.* **17**, 61–76.

1113. M. QUINZII (1992), *Increasing Returns and Efficiency*, Oxford University Press, Oxford, UK.

1114. H. RADEMACHER (1919), Uber partielle und totale differenzierbarkeit von funktionen mehrerer varieabeln und über die transformation der doppelintegrale, *Math. Ann.* **79**, 340–359.

1115. D. RALPH (1994), A chain rule for nonsmooth composite functions via minimization, *Bull. Austral. Math. Soc.* **49**, 129–137.

1116. D. RALPH (2002), A stable homotopy approach to horizontal linear complementarity problems, *Control Cybernet.* **31**, 575–599.

1117. D. RALPH AND S. J. WRIGHT (2004), Some properties of regularization and penalization schemes for MPECs, *Optim. Meth. Soft.* **5**, 527–556.

1118. F. RAMPAZZO AND R. B. VINTER (2000), Degenerate optimal control problems with state constraints, *SIAM J. Control Optim.* **39**, 989–1007.

1119. T. RAPCSÁK (1997), *Smooth Nonlinear Optimization in $I\!R^n$*, Kluwer, Dordrecht, The Netherlands.

1120. J.-P. Raymond (1997), Nonlinear boundary control of semilinear parabolic problems with pointwise state constraints, *Disc. Cont. Dynam. Systems* **3**, 341–370.

1121. J.-P. Raymond and H. Zidani (1998), Pontryagin's principle for state-constrained control problems governed by parabolic equations with unbounded controls, *SIAM J. Control Optim.* **38**, 1853–1879.

1122. J. Renegar (1995), Incorporating condition measures into the complexity theory of linear programming, *SIAM J. Optim.* **5**, 506–524.

1123. J. Renegar (1995), Linear programming, complexity theory and elementary functional analysis, *Math. Progr.* **70**, 279–540.

1124. L. Rifford (2002), Semiconcave control Lyapunov functions and stabilizing feedback, *SIAM J. Control Optim.* **41**, 659–681.

1125. S. M. Robinson (1972), Normed convex processes, *Trans. Amer. Math. Soc.* **174**, 127–140.

1126. S. M. Robinson (1973), Bounds for error in the solution set of a perturbed linear program, *Linear Algebra Appl.* **6**, 69–81.

1127. S. M. Robinson (1975), An application of error bounds for convex programming in a linear space, *SIAM J. Control* **13**, 271–273.

1128. S. M. Robinson (1976), Stability theory for systems of inequalities, II: Differentiable nonlinear systems, *SIAM J. Numer. Anal.* **13**, 130–143.

1129. S. M. Robinson (1976), Regularity and stability for convex multivalued functions, *Math. Oper. Res.* **1**, 130–143.

1130. S. M. Robinson (1979), Generalized equations and their solutions, I: Basic theory, *Math. Progr. Study* **10**, 128–141.

1131. S. M. Robinson (1980), Strongly regular generalized equations, *Math. Oper. Res.* **5**, 43–62.

1132. S. M. Robinson (1981), Some continuity properties of polyhedral multifunctions, *Math. Progr. Study* **14**, 206–214.

1133. S. M. Robinson (1982), Generalized equations and their solutions, II: Applications to nonlinear programming, *Math. Progr. Study* **19**, 200–221.

1134. S. M. Robinson (1983), Generalized equations, in *Mathematical Programming: The State of the Art*, edited by A. Bachem et al., pp. 346–367, Springer, Berlin.

1135. S. M. Robinson (1987), Local epi-continuity and local optimization, *Math. Progr. Study* **37**, 208–222.

1136. S. M. Robinson (1991), An implicit function theorem for a class of nonsmooth functions, *Math. Oper. Res.* **16**, 292–309.

1137. S. M. Robinson (2003), Constraint nondegeneracy in variational analysis, *Math. Oper. Res.* **28**, 201–232.

1138. S. M. Robinson (2003), Variational conditions with smooth constraints: Structure and analysis, *Math. Progr.* **97**, 245–265.

1139. S. M. Robinson (2004), Localized normal maps and the stability of variational conditions, *Set-Valued Anal.* **12**, 259–274.

1140. R. T. Rockafellar (1963), *Convex Functions and Dual Extremum Problems*, Ph.D. dissertation, Department of Mathematics, Harvard University, Cambridge, Massachusetts.

1141. R. T. Rockafellar (1966), Characterization of subdifferentials of convex functions, *Pacific J. Math.* **17**, 497–510.

1142. R. T. Rockafellar (1970), *Convex Analysis*, Princeton University Press, Princeton, New Jersey.

1143. R. T. Rockafellar (1970), Conjugate convex functions in optimal control and the calculus of variations, *J. Math. Anal. Appl.* **32**, 174–222.
1144. R. T. Rockafellar (1970), Generalized Hamiltonian equations for convex problems of Lagrange, *Pacific J. Math.* **33**, 411–427.
1145. R. T. Rockafellar (1971), Existence and duality theorems for convex problems of Bolza, *Trans. Amer. Math. Soc.* **159**, 1–39.
1146. R. T. Rockafellar (1979), Clarke's tangent cone and the boundaries of closed sets in $\mathbb{R}^n$, *Nonlinear Anal.* **3**, 145–154.
1147. R. T. Rockafellar (1979), Directional Lipschitzian functions and subdifferential calculus, *Proc. London Math. Soc.* **39**, 331–355.
1148. R. T. Rockafellar (1980), Generalized directional derivatives and subgradients of nonconvex functions, *Canad. J. Math.* **32**, 157–180.
1149. R. T. Rockafellar (1981), *The Theory of Subgradients and Its Applications to Problems of Optimization: Convex and Nonconvex Functions*, Helderman Verlag, Berlin.
1150. R. T. Rockafellar (1981), Proximal subgradients, marginal values and augmented Lagrangians in nonconvex optimization, *Math. Oper. Res.* **6**, 424–436.
1151. R. T. Rockafellar (1982), Favorable classes of Lipschitz continuous functions in subgradient optimization, in *Progress in Nondifferentiable Optimization*, edited by E. A. Nurminskii, pp. 125–143, IIASA, Laxenburg, Austria.
1152. R. T. Rockafellar (1982), Lagrange multipliers and subderivatives of optimal value functions in nonlinear programming, *Math. Progr. Study* **17**, 28–66.
1153. R. T. Rockafellar (1985), Maximal monotone relations and the second derivatives of nonsmooth functions, *Ann. Inst. H. Poincaré: Analyse Non Linéaire* **2**, 167–184.
1154. R. T. Rockafellar (1985), Lipschitzian properties of multifunctions, *Nonlinear Anal.* **9**, 867–885.
1155. R. T. Rockafellar (1985), Extensions of subgradient calculus with applications to optimization, *Nonlinear Anal.* **9**, 665–698.
1156. R. T. Rockafellar (1988), First- and second-order epi-differentiability in nonlinear programming, *Trans. Amer. Math. Soc.* **307**, 75–108.
1157. R. T. Rockafellar (1989), Letter to B. S. Mordukhovich of March 6, 1989, Seattle, Washington.
1158. R. T. Rockafellar (1989), Derivation of some improved formulas in subdifferential calculus, privately circulated note.
1159. R. T. Rockafellar (1989), Proto-differentiability of set-valued mappings and its applications in optimization, in *Analyse Non Linéaire*, edited by H. Attouch et al., pp. 449–482, Gauthier-Villars, Paris.
1160. R. T. Rockafellar (1993), Lagrange multipliers and optimality, *SIAM Rev.* **35**, 183–238.
1161. R. T. Rockafellar (1993), Dualization of subgradient conditions for optimality, *Nonlinear Anal.* **20**, 627–642.
1162. R. T. Rockafellar (1996), Equivalent subgradient versions of Hamiltonian and Euler-Lagrange equations in variational analysis, *SIAM J. Control Optim.* **34**, 1300–1315.
1163. R. T. Rockafellar (2000), Second-order convex analysis, *J. Nonlinear Convex Anal.* **1**, 1–16.
1164. R. T. Rockafellar (2004), Hamilton–Jacobi theory and parametric analysis in fully convex problems of optimal control, *J. Global Optim.* **28**, 419–431.

1165. R. T. ROCKAFELLAR AND R. J-B. WETS (1998), *Variational Analysis*, Springer, Berlin.
1166. R. T. ROCKAFELLAR AND P. R. WOLENSKI (2000), Convexity in Hamilton-Jacobi theory, I: Dynamics and duality, *SIAM J. Control Optim.* **39**, 1323–1350.
1167. R. T. ROCKAFELLAR AND P. R. WOLENSKI (2000), Convexity in Hamilton-Jacobi theory, II: Envelope representations, *SIAM J. Control Optim.* **39**, 1351–1372.
1168. R. T. ROCKAFELAR AND D. ZAGRODNY (1997), A derivative-coderivative inclusion in second-order nonsmooth analysis, *Set-Valued Anal.* **5**, 1–17.
1169. S. ROLEWICZ (1979), On paraconvex multifunctions, *Oper. Res. Verfahren* **31**, 540–546.
1170. S. ROLEWICZ (1979), On γ-paraconvex multifunctions, *Math. Japon.* **24**, 293–300.
1171. S. ROLEWICZ (2005), On differentiability of strongly $\alpha(\cdot)$-paraconvex functions in non-separable Asplund spaces, *Studia Math.*, **167**, 235–244.
1172. S. ROLEWICZ (2007), Paraconvex analysis on $C_{\mathbf{E}}^{1,u}$ manifolds, *Optimization*, **56**, 49–60.
1173. J. F. ROSENBLUETH AND R. B. VINTER (1991), Relaxation procedures for time delay systems, *J. Math. Anal. Appl.* **162**, 542–563.
1174. J. F. ROSENBLUETH, J. WARGA AND Q. J. ZHU (1997), On the characterization of properly relaxed delayed controls, *J. Math. Anal. Appl.* **209**, 274–290.
1175. I. M. ROSS AND F. FAHROO (2004), Legendre pseudospectral approximations of optimal control problems, *Lecture Notes Cont. Inf. Sci.* **295**, pp. 327–342, Springer, New York.
1176. R. ROSSI AND G. SAVARÉ (2005), Gradient flows of nonconvex functionals in Hilbert spaces and applications, *ESAIM: Control Optim. Cal. Var.*, to appear.
1177. T. ROUBIČEK (1997), *Relaxation in Optimization Theory and Variational Calculus*, De Gruyter, Berlin.
1178. J. D. L. ROWLAND AND R. B. VINTER (1992), Pontryagin type conditions for differential inclusions, *J. Math. Anal. Appl.* **165**, 587–597.
1179. J. D. L. ROWLAND AND R. B. VINTER (1993), Dynamic optimization problems with free-time and active state constraints, *SIAM J. Control Optim.* **31**, 677–697.
1180. L. I. ROZONOÉR (1959), L. S. Pontryagin maximum principle in the theory of optimal systems, I–III, *Autom. Remote Control* **20**, 1288–1302, 1405–1421, 1517–1532.
1181. A. M. RUBINOV (2000), *Abstract Convexity and Global Optimization*, Kluwer, Dordrecht, The Netherlands.
1182. A. M. RUBINOV (2002), Equilibrium with restriction on exchange, *Cybernetics and Systems Analysis*, No. 2, 55–70.
1183. J.-J. RÜCKMANN (1999), On existence and uniqueness of stationary points in semi-infinite programming, *Math. Progr.* **86**, 387–415.
1184. D. L. RUSSELL (1966), Optimal regulation of linear symmetruc hyperbolic systems with finite dimensional control, *SIAM J. Control* **4**, 276–294.
1185. L. D. SABBAGH (1969), Variational problems with lags, *J. Optim. Theory Appl.* **3**, 34–51.
1186. S. SAKS (1937), *Theory of the Integral*, 2nd edition, Hafner Publishing Co., New York.

1187. D. SALAMON (1984), *Control and Observation of Neutral Systems*, Pitman, Harlow, Essex, UK.
1188. P. A. SAMUELSON (1947), *Foundations of Economic Analysis*, Harvard University Press, Cambridge, Massachusetts.
1189. P. A. SAMUELSON (1954), The pure theory of public expenditures, *Rev. Econ. Stat.* **36**, 387–389.
1190. J. SCHAUDER (1930), Über die umkerhrung linearer, stetiger funktionaloperationen, *Studia Math.* **2**, 1–6.
1191. H. SCHEEL AND S. SCHOLTES (2000), Mathematical programs with equilibrium constraints: Stationarity, optimality and sensitivity, *Math. Oper. Res.* **25**, 1–22.
1192. S. SCHOLTES (2001), Convergence properties of a regularization scheme for mathematical programs with complementarity constraints, *SIAM J. Optim.* **11**, 818–936.
1193. S. SCHOLTES (2002), On the existence and computation of EPEC solutions, talk given at the ICCP Conference in Cambridge, UK.
1194. S. SCHOLTES AND M. STÖHR (1999), Exact penalization of mathematical programs with equilibrium constraints, *SIAM J. Control Optim.* **37**, 617–652.
1195. L. SCRIMALI (2004), Variational inequalities and optimal equilibrium distributions in transportation networks, *Math. Ineq. Appl.* **7**, 439–451.
1196. R. SCHULTZ (2000), Some aspects of stability in stochastic programming, *Ann. Oper. Res.* **100**, 55–84.
1197. L. SCHWARTZ (1966), *Théorie des Distributions*, Hermann, Paris.
1198. A. SEEGER (1992), Limiting behavior of the approximate second-order subdifferential of a convex function, *J. Optim. Theory Appl.* **74**, 527–544.
1199. A. SEEGER (1994), Approximate Euler-Lagrange inclusion, approximate transversality condition, and sensitivity analysis of convex parametric problems of calculus of variations, *Set-Valued Anal.* **2**, 307–325.
1200. H. S. SENDOV (2000), *Variational Spectral Analysis*, Ph.D. dissertation, Department of Combinatorics and Optimization, University of Waterloo, http://etd/uwaterloo.ca/etd/hssendov2000.pdf, Waterloo, Canada.
1201. B. SENDOV AND V. A. POPOV (1988), *The Averaged Moduli of Smoothness*, Wiley, New York.
1202. F. SEVERI (1930), Su alcune questioni di topologia infinitesimale, *Ann. Soc. Polon. Math.* **9**, 97–108.
1203. A. SHAPIRO (1988) Sensitivity analysis of nonlinear programs and differentiability property of metric projections, *SIAM J. Control Optim.* **26**, 628–645.
1204. A. SHAPIRO (1990), On concepts of directional defferentiability, *J. Math. Anal. Appl.* **66**, 477–487.
1205. A. SHAPIRO (1994), Sensitivity analysis of parameterized programs via generalized equations, *SIAM J. Control Optim.* **32**, 553–571.
1206. A. SHAPIRO, T. HOMEM-DE-MELLO AND J. KIM (2002), Conditioning of convex pievewise linear stochastic programs, *Math. Progr.* **94**, 1–19.
1207. N. Z. SHOR (1972), On a class of almost-differentiable functions and on a minimization method for functions from this class, *Kibernetika*, No. 4. 65–70.
1208. N. Z. SHOR (1985), *Minimization Methods for Non-Differentiable Functions*, Springer, Berlin.
1209. I. A. SHVARTSMAN (2007), New approximation method in the proof of the maximum principle for nonsmooth optimal control problems with state constraints, *J. Math. Anal. Appl.*, **326**, 974-1000.

1210. I. A. SHVARTSMAN AND R. B. VINTER (2006), Regularity properties of optimal controls for state constrained problems with time-varying control constraints, *Nonlinear Anal.*, **65**, 448-474.
1211. G. H. SILVA AND R. B. VINTER (1998), Necessary conditions for optimal impulsive control problems, *SIAM J. Control Optim.* **35**, 1829–1846.
1212. J. SIMON (1987), Compact sets in the space $L^p(0,T;B)$, *Ann Mat. Pura Appl.* **146**, 65–96.
1213. S. SIMONS (1998), *Minimax and Monotonicity*, Springer, Berlin.
1214. I. SINGER (1997), *Abstract Convex Analysis*, Wiley, New York.
1215. G. V. SMIRNOV (1991), Discrete approximations and optimal solutions to differential inclusions, *Cybernetics* **27**, 101–107.
1216. G. V. SMIRNOV (2001), *Introduction to the Theory of Differential Inclusions*, American Mathematical Society, Providence, Rhode Island.
1217. A. SMITH (1776), *An Inquiry into the Nature and Causes of the Wealth of Nations*, Clarendon Press, Oxford, England.
1218. S. L. SOBOLEV (1963), *Applications of Functional Analysis in Mathematical Physics*, American Mathematical Society, Providence, Rhode Island.
1219. M. V. SOLODOV AND B. F. SVAITER (2000), Error bounds for proximal point subproblems and associated inexasct proximal point algorithms, *Math. Progr.* **88**, 371–389.
1220. E. SONTAG (1999), Stability and stabilization: Discontinuity and the effect of disturbances, in *Nonlinear Analysis, Differential Equations, and Control*, edited by F. H. Clarke and R. J. Stern, pp. 551–598, Kluwer, Dordrecht, The Netherlands.
1221. V. A. SROCHKO (1984), Optimality conditions of the maximum principle type in Goursat-Darboux systems, *Siberian J. Math.* **25**, 126–132.
1222. H. VON STACKELBERG (1934), *Marktform und Gleichgewicht*, Springer, Berlin.
1223. G. STAMPACCHIA (1964), Formes bilinéaires coercitives sur les ensemble convexes, *C. R. Acad. Sci. Paris* **258**, 4413–4416.
1224. C. STEGALL (1978), Optimization of functions on certain subsets of Banach spaces, *Trans. Amer. Math. Soc.* **236**, 171–176.
1225. O. STEIN (2003), *Bilevel Strategies in Semi-Infinite Programming*, Kluwer, Boston, Massachusetts.
1226. A. S. STREKALOVSKY (1987), On the problem of global extremum, *Soviet Math. Dokl.* **35**, 194–198.
1227. A. S. STREKALOVSKY (1998), Global optimality conditions for nonconvex optimization, *J. Global Optim.* **12**, 415–434.
1228. A. S. STREKALOVSKY (2003), *Elements of Nonconvex Optimization*, Nauka, Novosibirsk.
1229. M. STUDNIARSKI AND D. E. WARD (1999), Weak sharp minima: Characterizations and sufficient conditions, *SIAM J. Control Optim.* **38** (1999), 219–236.
1230. A. I. SUBBOTIN (1995), *Generalized Solutions of First-Order PDEs*, Birkhäuser, Boston, Massuchusetts.
1231. N. N. SUBBOTINA (1989), The maximum principle and the superdifferential of the value function, *Prob. Control Inform. Theory* **18**, 151–160.
1232. F. SULLIVAN (1981), A characterization of complete metric spaces, *Proc. Amer. Math. Soc.* **83**, 345–346.
1233. M. I. SUMIN (2000), Optimal control of semilinear elliptic equations with state constraints: Maximum principle for minimizing sequences, regularity, normality, sensitivity, *Control Cybernet.* **29**, 449–472.

1234. M. I. SUMIN (2001), Suboptimal control of a semilinear elliptic equation with a phase constraint and a boundary control, *Diff. Eq.* **37**, 281–300.
1235. H. J. SUSSMANN (1994), A strong version of the Lojasiewicz maximum principle, in *Optimal Control and Differential Equations*, edited by N. N. Pavel, pp. 293–309, Marcel Dekker, New York.
1236. H. J. SUSSMANN (1998), Geometry and optimal control, in *Mathematical Control Theory*, edited by J. Baillieul and J. C. Willems, pp. 140-198, Springer, New York.
1237. H. J. SUSSMANN (2000), New theories of set-valued differentials and new versions of the maximum principle, in *Nonlinear Control in the Year 2000*, edited by A. Isidori, F. Lamnabhi-Lagarrigue and W. Respondek, pp. 487–526, Springer, Berlin.
1238. H. J. SUSSMANN (2002), Needle variations and almost lower semicontinuous differential inclusions, *Set-Valued Anal.* **10**, 233–285.
1239. H. J. SUSSMANN AND J. V. WILLEMS (1997), 300 years of optimal control: From the brachystochrone to the maximum principle, *IEEE Control Syst. Magaz.* **17**, 32–44.
1240. A. ŚWIECH (1994), Unbounded second order partial differential equations in infinite dimensional Hilbert spaces, *Comm. Part. Diff. Eq.* **19**, 1999-2036.
1241. A. ŚWIECH (1996), Sub- and superoptimality principles of dynamic programming revisited, *Nonlinear Anal.* **26**, 1429–1436.
1242. T. TADUMADZE AND L. ALKHAZISHVILI (2003), Formulas of variations of solutions for nonlinear controlled delay differential equations with discontinuous initial conditions, *Mem. Diff. Eq. Math. Phys.* **29**, 125–150.
1243. W. TAKAHASHI (2000), *Nonlinear Functional Analysis*, Yokohama Publishers, Yokohama, Japan.
1244. L. THIBAULT (1976), Probléme de Bolza dans un espace de Banach séparable, *C. R. Acad. Sci. Paris* **282**, 1303–1306.
1245. L. THIBAULT (1978), Sous-différentiel de fonctions vertorielles compactement lipschitziennes, *C. R. Acad. Sci. Paris* **286**, 995–998.
1246. L. THIBAULT (1980), Subdifferentials of compactly Lipschitzian vector functions, *Ann. Mat. Pura Appl.* **125**, 157–192.
1247. L. THIBAULT (1982), On generalized differentials and subdifferentials of Lipschitz vector-valued functions, *Nonlinear Anal.* **6**, 1037–1053.
1248. L. THIBAULT (1983), Tangent cones and quasi-interiorly tangent cones to multifunctions, *Trans. Amer. Math. Soc.* **277**, 601–621.
1249. L. THIBAULT (1991), On subdifferentials of optimal value functions, *SIAM J. Control Optim.* **29**, 1019–1036.
1250. L. THIBAULT (1992), Lagrange-Kuhn-Tucker multipliers for general mathematical programming problems, in *Optimization and Nonlinear Analysis*, edited by A. Ioffe, L. Marcus and S. Reich, Pitman Researh Notes Math. Ser. **244**, pp. 311–315, Longman, Harlow, Essex, UK.
1251. L. THIBAULT (1995), A note on the Zagrodny mean value theorem, *Optimization* **35**, 127–130.
1252. L. THIBAULT (1997), On compactly Lipschitzian mappings, in *Recent Advances in Optimization*, edited by P. Gritzmann et al., Lecture Notes Econ. Math. Syst. **456**, pp. 356–364, Springer, Berlin.
1253. L. THIBAULT (2003), Sweeping process with regular and nonregular sets, *J. Diff. Eq.* **193**, 1–26.

1254. L. THIBAULT AND D. ZAGRODNY (1995), Integration of subdifferentials of lower semicontinuous functions, *J. Math. Anal. Appl.* **189**, 22–58.

1255. D. TIBA (1990), *Optimal Control of Nonsmooth Distributed Parameter Systems*, Springer, Berlin.

1256. V. M. TIKHOMIROV (1990), *Stories about Maxima and Minima*, American Mathematical Society, Providence, Rhode Island.

1257. V. M. TIKHOMIROV (1997), *Elements of the Theory of Extrema*, Tinbergen Institute, TI 97-048/4, Amsterdam, The Netherlands.

1258. A. A. TOLSTONOGOV (2000), *Differential Inclusions in a Banach Space*, Kluwer, Dordrecht, The Netherlands.

1259. A. A. TOLSTONOGOV (2004), Properties of the attainable sets of evolution inclusions and control systems of subdifferential type, *Siberian Math. J.* **45**, 763–784.

1260. L. TONELLI (1923), *Fondamenti di Calcolo delle Variazioni*, published in two volumes, Zanichelli, Bologna, Italy.

1261. R. TOURKY (1998), A new approach to the limit theorem on the core of an economy in vector lattices, *J. Econ. Theory* **78**, 321–328.

1262. J. S. TREIMAN (1983), Characterization of Clarke's tangent and normal cones in finite and infinite dimensions, *Nonlinear Anal.* **7**, 771–783.

1263. J. S. TREIMAN (1986), Clarke's gradients and epsilon-subgradients in Banach spaces, *Trans. Amer. Math. Soc.* **294**, 65–78.

1264. J. S. TREIMAN (1988), Shrinking generalized gradients, *Nonlinear Anal.* **12**, 1429–1450.

1265. J. S. TREIMAN (1989), Finite dimensional optimality conditions: B-gradients, *J. Optim. Theory Appl.* **62**, 771–783.

1266. J. S. TREIMAN (1990), Optimal control with small generalized gradients, *SIAM J. Control Optim.* **28**, 720–732.

1267. J. S. TREIMAN (1995), The linear nonconvex generalized gradients and Lagrange multipliers, *SIAM J. Optim.* **5**, 670–680.

1268. J. S. TREIMAN (1999), Lagrange multipliers for nonconvex generalized gradients with equality, inequality, and set constraints, *SIAM J. Control Optim.* **37**, 1313–1329.

1269. J. S. TREIMAN (2002), The linear generalized gradient in infinite dimensions, *Nonlinear Anal.* **48**, 427–443.

1270. A. A. TRET'YAKOV AND J. E. MARSDEN (2003), Factor-analysis of nonlinear mappings: p-Regularity theory, *Commun. Pure Appl. Anal.* **2**, 425–445.

1271. F. TRÖLTZSCH (1984), *Optimality Conditions for Parabolic Control Problems and Applications*, Teubner Texte, Leipzig.

1272. I. TSEVENDORJ (2001), Piecewise-convex maximization problems, *J. Global Optim.* **21**, 1–14.

1273. H. D. TUAN (1995), On controllability and extremality in nonconvex differential inclusions, *J. Optim. Theory Appl.* **85**, 435–472.

1274. A. UDERZO (2002), Notes on metric regularity and open covering, Technical Report MS2.7.8, pp. 1–26, Dipartimento di Sistemi ed Istituzioni per l'Economia, Univ. di L'Aquila.

1275. C. URSESCU (1975), Multifunctions with closed convex graphs, *Czech. Math. J.* **25**, 438–441.

1276. C. URSESCU (1982), Tangent sets' calculus and necessary conditions for extremality, *SIAM J. Control Optim.* **20**, 563–574.

1277. M. VALADIER (1990), Young measures, in *Methods of Convex Analysis*, Lecture Notes Math. **1446**, pp. 152–188, Springer, Berlin.
1278. P. P. VARAJAY (1967), Nonlinear programming in Banach spaces, *SIAM J. Appl. Math.* **15**, 284–293.
1279. F. P. VASILIEV (1969), Optimality conditions for certain classes of systems which are not solved with respect to the derivative, *Soviet Math. Dokl.* **10**, 224–227.
1280. F. P. VASILIEV (1981), *Solution Methods for Extremal Problems*, Nauka, Moscow.
1281. O. V. VASILIEV (1996), *Optimization Methods*, World Federation Publishers, Atlanta, Georgia.
1282. V. VELIOV (1994), Differential inclusions with stable subinclusions, *Nonlinear Anal.* **23**, 1027–1038.
1283. V. VELIOV (1997), Lipschitz continuity of the value function in optimal control, *J. Optim. Theory Appl.* **94**, 335–363.
1284. V. VELIOV (1997), On the time-discretization of control systems, *SIAM J. Control Optim.* **35**, 1470–1486.
1285. I. Y. VERCHENKO AND A. N. KOLMOGOROV (1934), Continuation of investigation on discontinuity points for functions of two variables, *Dokl. Acad. Nauk SSSR* **4**, 361–364.
1286. J.-P. VIAL (1983), Strong and weak convexity of sets and functions, *Math. Oper. Res.* **8**, 231–259.
1287. A. VILLAR (2000), *Equilibrium and Efficiency in Production Economies*, Springer, Berlin.
1288. A. VILLAR (2001), On the efficiency of market equilibrium in production economies, *J. Global Optim.* **20**, 375–389.
1289. R. B. VINTER (2000), *Optimal Control*, Birkhäuser, Boston, Massachusetts.
1290. R. B. VINTER AND G. PAPPAS (1982), A maximum principle for non-smooth optimal control problems with state constraints, *J. Math. Anal. Appl.* **89**, 212–232.
1291. R. B. VINTER AND F. L. PEREIRA (1988), A maximum principle for optimal processes with discontinuous trajectories, *SIAM J. Control Optim.* **26**, 205–229.
1292. R. B. VINTER AND P. R. WOLENSKI (1990), Coextremals and the value function for control problems with data measurable in time, *J. Optim. Theory Appl.* **153**, 37–51.
1293. R. B. VINTER AND P. D. WOODFORD (1997), On the occurrence of intermediate local minimizers that are not strong local minimizers, *Systems Cont. Lett.* **31**, 235–342.
1294. R. B. VINTER AND H. ZHENG (1997), The extended Euler-Lagrange condition for nonconvex variational problems, *SIAM J. Control Optim.* **35**, 56–77.
1295. R. B. VINTER AND H. ZHENG (1998), The extended Euler-Lagrange condition for nonconvex variational problems with state constraints, *Trans. Amer. Math. Soc.* **350**, 1181–1204.
1296. R. B. VINTER AND H. ZHENG (2000), Necessary conditions for free end-time, measurable time dependent optimal control problems with state constraints, *Set-Valued Anal.* **8**, 10–29.
1297. R. B. VINTER AND H. ZHENG (2003), Some finance problems solved with nonsmooth optimization techniques, *J. Optim. Theory Appl.* **119**, 1–18.

1298. V. Volterra (1931), *Théories Mathématoque de la Lutte pour la Vie*, Guathoer-Villar, Paris.
1299. D. W. Walkup and R. J-B. Wets (1969), A Lipschitzian characterization of convex polyhedra, *Proc. Amer. Math. Soc.* **20**, 167–173.
1300. L. Walras (1874-7), *Eléments d'Economie Politique Pure*, L. Corbaz and Company, Lausanne.
1301. L. Wang (2005), Discrete approximations to optimization of neutral functional differential inclusions, *J. Math. Anal. Appl.* **309**, 474–488.
1302. P. K. C. Wang (1964), Control of distributed parameter systems, in *Advances in Control Systems, Theory and Applications*, edited by C. T. Leondes, Vol. 1, pp. 75–172, Academic Press, New York.
1303. X. Wang (2005), Subdifferentiability of real functions, *Real Anal. Exchange* **30**, 137–171.
1304. X. Wang (2007), Extremal characterizations of reflexive spaces, *J. Aust. Math. Soc.* **82**, 429–439.
1305. X. Wang and V. Jeyakumar (2000), A sharp Lagrange multiplier rule for nonsmooth mathematical programming problems involving equality constraints, *SIAM J. Optim.* **10**, 1136–1148.
1306. D. E. Ward (1993), Calculus for parabolic second-order derivatives, *Set-Valued Anal.* **1**, 213–246.
1307. D. E. Ward (1994), A chain rule for parabolic second-order epiderivatives, *Optimization* **28**, 223–236.
1308. D. E. Ward (1995), A comparison of second-order epiderivatives: Calculus and optimality conditions, *J. Math. Anal. Appl.* **193**, 465–482.
1309. D. E. Ward (1996), Dini derivatives of the marginal functions of a non-Lipschitzian program, *SIAM J. Optim.* **6**, 198–211.
1310. D. E. Ward (1999), Second-order necessary conditions in nonsmooth programming, in *Systems Modelling and Optimization*, edited by M. P. Polis et al., pp. 216–224, Chapman and Hall/CRC Press, Boca Raton, Florida.
1311. D. E. Ward and J. M. Borwein (1987), Nonconvex calculus in finite dimensions, *SIAM J. Control Optim.* **25**, 1312–1340.
1312. D. E. Ward and G. M. Lee (2001), Generalized properly efficient solutions of vector optimization problems, *Math. Oper. Res.* **53**, 215–232.
1313. J. Warga (1962), Relaxed variational problems, *J. Math. Anal. Appl.* **4**, 111–128.
1314. J. Warga (1962), Necessary conditions for minimum in relaxed variational problems, *J. Math. Anal. Appl.* **4**, 129–145.
1315. J. Warga (1972), *Optimal Control of Differential and Functional Equations*, Academic Press, New York.
1316. J. Warga (1975), Necessary optimality conditions without differentiability assumptions in optimal control, *J. Diff. Eq.* **15**, 41–62.
1317. J. Warga (1976), Necessary conditions withiut differentiability assumptions in unilatateral control problems, *J. Diff. Eq.* **21**, 25–38.
1318. J. Warga (1976), Derivate containers, inverse functions, and controllability, in *Calculus of Variations and Control Theory*, edited by D. L. Russel, pp. 13–46, Academic Press, New York.
1319. J. Warga (1978), Controllability and a multiplier rule for nondifferentiable optimization problems, *SIAM J. Control Optim.* **16**, 803–812.
1320. J. Warga (1981), Fat homeomorphisms and unbounded derivate containers, *J. Math. Anal. Appl.* **81**, 545–560.

1321. J. WARGA (1983), Controllability, extremality and abnormality in nonsmooth optimal control, *J. Optim. Theory Appl.* **41**, 239–260.
1322. J. WARGA (1988), Homeomorphisms and local $\mathcal{C}^1$ approximations, *Nonlinear Anal.* **12**, 593–597.
1323. J. WARGA (1995), A proper relaxation of control with variable shifts, *J. Math. Anal. Appl.* **196**, 783–793.
1324. D. WASHBURN (1979), A bound on the boundary input map for parabolic equations with applications to time optimal control, *SIAM J. Control Optim.* **17**, 652–671.
1325. T. WAŻEWSKI (1961), Systémes de commandes et équations au contingent, *Bull. Acad. Pol. Sci., Ser. Math. Ast. Phys.* **9**, 151–155.
1326. K. WEIERSTRASS (1927), *Vorlesungen über Variationsrechnung*, Akademische Verlagsgesellschaft, Leipzig.
1327. R. J-B. WETS (1996), Challenges in stochastic programming, *Math. Progr.* **75**, 115–135.
1328. L. W. WHITE (1985), Distributed control of a hyperbolic problem with control and stress constraints, *J. Math. Anal. Appl.* **106**, 41–53.
1329. P. R. WOLENSKI (1990), The exponential formula for the reachable set of a Lipschitz differential inclusion, *SIAM J. Control Optim.* **28**, 1148–1161.
1330. P. R. WOLENSKI AND Y. ZHUANG (1998), Proximal analysis and the minimal time function, *SIAM J. Control Optim.* **36**, 1048–1072.
1331. P. D. WOODFORD (1997), *Optimal Control for Nonsmooth Dynamic Systems*, Ph.D. dissertation, Department of Electric. Electron. Eng., Imperial College, London, UK.
1332. M. H. WRIGHT (1999), Ill-conditioning and computational error in interior methods for nonlinear programming, *SIAM J. Optim.* **9**, 84–111.
1333. S. E. WRIGHT (1995), Consistency of primal-dual approximations for convex optimal control, *SIAM J. Control Optim.* **33**, 1489–1509.
1334. Z. WU AND J. J. YE (2002), On error bounds for lower semicontinuous functions, *Math. Progr.* **92**, 421–435.
1335. Z. WU AND J. J. YE (2003), First-order and second-order conditions for error bounds, *SIAM J. Optim.* **14**, 621–645.
1336. Z. WU AND J. J. YE (2003), Equivalence between various derivatives and subdifferentials of the distance function, *J. Math. Anal. Appl.* **282**, 629–647.
1337. X. Q. YANG AND V. JEYAKUMAR (1992), Generalized second-order directional derivatives and optimization with $C^{1,1}$ functions, *Optimization* **26**, 165–185.
1338. J. J. YE (1999), Optimality conditions for optimization problems with complementarity constraints, *SIAM J. Control Optim.* **9**, 374–387.
1339. J. J. YE (2000), Constraint qualifications and necessary optimality conditions for optimization problems with variational inequality constraints, *SIAM J. Optim.* **10**, 943–962.
1340. J. J. YE (2001), Multiplier rules under mixed assumptions on differentiability and Lipschitz continuity, *SIAM J. Control Optim.* **39**, 1441–1460.
1341. J. J. YE (2004), Nondifferentiable multiplier rules for optimization and bilevel optimization problems, *SIAM J. Optim.* **15**, 252–274.
1342. J. J. YE (2005), Necessary and sufficient optimality conditions for mathematical programs with equilibrium constraints, *J. Math. Anal. Appl.* **307**, 350–369.

1343. J. J. YE AND X. Y. YE (1997), Necessary optimality conditions for optimization problems with variational inequality constraints, *Math. Oper. Res.* **22**, 977–997.

1344. J. J. YE, D. L. ZHU AND Q. J. ZHU (1997), Exact penalization and necessary optimality conditions for generalized bilevel programming problems, *SIAM J. Optim.* **7**, 481–507.

1345. J. J. YE AND Q. J. ZHU (2003), Multiobjective optimization problems with variational inequality constraints, *Math. Progr.* **96**, 139–160.

1346. N. D. YEN (1995), Lipschitz continuity of solutions of variational inequalities with a parametric polyhedral constraint, *Math. Oper. Res.* **20**, 695–708.

1347. N. D. YEN (1995), Hölder continuity of solutions to a parametric variational inequality, *Appl. Math. Optim.* **31**, 245–255.

1348. D. YOST (1993), Asplund spaces for beginners, *Acta Univ. Carolinae, Ser. Math. Phys.* **34**, 159–177.

1349. L. C. YOUNG (1933), On approximation by polygons in the calculus of variations, *Proc. Royal Soc.* (A) **141**, 325–341.

1350. L. C. YOUNG (1937), Generalized curves and the existence of an attained absolute minimum in the calculus of variations, *C. R. Soc. Sci. Lett. Varsovie*, Cl. III, **30**, 212–234.

1351. L. C. YOUNG (1969), *Lectures on the Calculus of Variations and Optimal Control Theory*, Saunders, Philadelphia, Pennsylvania.

1352. D. ZAGRODNY (1988), Approximate mean value theorem for upper subderivatives, *Nonlinear Anal.* **12**, 1413–1428.

1353. C. ZĂLINESCU (2002), *Convex Analysis in General Vector Spaces*, World Scientific, Singapore.

1354. W. I. ZANGWILL (1967), Nonlinear programming via penalty functions, *Management Sci.* **13**, 344–358.

1355. S. C. ZAREMBA (1936), Sur les équations au paratingent, *Bull. Sci. Math.* **60**, 139–160.

1356. A. J. ZASLAVSKI (2000), Generic well-posedness of optimal control problems without convexity assumptions, *SIAM J. Control Optim.* **39**, 250–280.

1357. A. J. ZASLAVSKI (2006), *Turnpike Properties in the Calculus of Variations and Optimal Control*, Springer, New York.

1358. V. M. ZEIDAN (2001), New second-order optimality conditions for variational problems with C^2-Hamiltonians, *SIAM J. Control Optim.* **40**, 577–609.

1359. M. I. ZELIKIN AND N. B. MELNIKOV (2004), The Bellman function and optimal synthesis in control problems with nonsmooth constraints, *J. Math. Sci.* **121**, 2281–2294.

1360. R. ZHANG (1994), Problems of hierarchical optimization in finite dimensions, *SIAM J. Optim.* **4**, 521–536.

1361. R. ZHANG (2003), Multistage bilevel programming problems, *Optimization* **52**, 605–616.

1362. R. ZHANG (2005), Weakly upper Lipschitzian multifunctions and applications to parametric optimization, *Math. Progr.* **102**, 153–166.

1363. R. ZHANG AND J. S. TREIMAN (1995), Upper-Lipschitz multifunctions and inverse subdifferentials, *Nonlinear Anal.* **24**, 273–286.

1364. X. Y. ZHENG AND K. F. NG (2005), The Fermat rule for multifunctions in Banach spaces, *Math. Progr.* **104**, 69–90.

1365. X. Y. Zheng and K. F. Ng (2007), Metric subregularity and constraint qualifications for convex generalized equations in Banach spaces, *SIAM J. Optim.* **18**, 437–460.
1366. X. Y. Zhou (1990), Maximum principle, dynamic programming and their connections in deterministic control, *J. Optim. Theory Appl.* **65**, 363–373.
1367. X. Y. Zhou (1995), Deterministic near-optimal controls, part I: Necessary and sufficient conditions for near optimality, *J. Optim. Theory Appl.* **85**, 473–488.
1368. X. Y. Zhou (1996), Deterministic near-optimal controls, part II: Dynamic programming and viscosity solution approach, *Math. Oper. Res.* **21**, 655–674.
1369. X. Y. Zhou (1998), Stochastic near-optimal controls: Necessary and sufficient conditions for near-optimality, *SIAM J. Control Optim.* **36**, 929–947.
1370. Q. J. Zhu (1996), Necessary optimality conditions for nonconvex differential inclusions with endpoint constraints, *J. Diff. Eq.* **124**, 186–204.
1371. Q. J. Zhu (1998), The equivalence of several basic theorems for subdifferentials, *Set-Valued Anal.* **6**, 171–185.
1372. Q. J. Zhu (2000), Hamiltonian necessary conditions for a multiobjective optimal control problem, *SIAM J. Control Optim.* **39**, 97–112.
1373. Q. J. Zhu (2002), Necessary conditions for optimization problems in smooth Banach spaces, *SIAM J. Optim.* **12**, 1032–1047.
1374. Q. J. Zhu (2003), Lower semicontinuous Lyapunov functions and stability, *J. Nonlinear Convex Anal.* **4**, 325–332.
1375. Q. J. Zhu (2004), Nonconvex separation theorem for multifunctions, subdifferential calculus and applications, *Set-Valued Anal.* **12**, 275–290.
1376. T. Zolezzi (2002), On the distance theorem in quadratic optimization, *J. Convex Anal.* **9**, 693–700.
1377. T. Zolezzi (2003), Condition number theorems in optimization, *SIAM J. Optim.* **14**, 507–516.
1378. J. Zowe and S. Kurcyusz (1979), Regularity and stability for the mathematical programming problem in Banach spaces, *Appl. Math. Optim.* **5**, 49–62.
1379. E. Zuazua (2005), Propagation, observation, and control of waves approximated by finite difference methods, *SIAM Rev.* **47**, 197–243.

陈　述　表

第 5 章

第 6 章

第 7 章

第 8 章

记　号　表

运算和记号

$:=, =:$　定义为

$\equiv$　恒等于

$*$　表示某些对偶/伴随/极化运算

$\langle\cdot,\cdot\rangle$　空间 X 和它的拓扑对偶 X^* 之间的典范偶对

$x \to \bar{x}$　x 强 (范数) 收敛于 $\bar{x}$

$x \xrightarrow{w} \bar{x}$　x 弱 *(在弱 * 拓扑下) 收敛于 $\bar{x}$

$x \xrightarrow{w^*} \bar{x}$　x 弱 (在弱拓扑下) 收敛于 $\bar{x}$

$x \xrightarrow{\Omega} \bar{x}$　x 收敛于 $\bar{x}$ 且 $x \in \Omega$

$\liminf$　实数的下极限

$\limsup$　实数的上极限

Liminf　集值映射的下/内极限

Limsup　集值映射的上/外极限

$\dim X$, $\mathrm{codim} X$　X 的维数和余维数

$\prec$　序关系

$\|\cdot\|, |\cdot|, |||\cdot|||$　范数

$\mathrm{haus}(\Omega_1, \Omega_2)$　集合之间的 Pompieu-Hausdorff 距离

$\mathrm{lip}F(\bar{x},\bar{y})$　F 在 $(\bar{x},\bar{y})$ 附近的确切 Lipschitz 界

$\mathrm{reg}F(\bar{x},\bar{y})$　F 在 $(\bar{x},\bar{y})$ 附近的确切度量正则界

$\mathrm{cov}F(\bar{x},\bar{y})$　F 在 $(\bar{x},\bar{y})$ 附近的确切覆盖/线性开性界

$\mathrm{rad}F(\bar{x},\bar{y})$　F 在 $(\bar{x},\bar{y})$ 附近度量正则性的半径

$\triangle$　证明结束

空间

$\mathbb{R} := (-\infty, \infty)$　实直线

$\overline{\mathbb{R}} := [-\infty, \infty]$　增广实直线

$\mathbb{R}^n$ n 维 Euclid 空间

$\mathbb{R}^n_+$, $\mathbb{R}^n_-$　$\mathbb{R}$ 的非负和非正象限

$\mathcal{C}([a,b];X)$　取上确界范数的 $[a,b]$ 到 X 的连续映射空间

$\mathcal{C}(K)$　紧集 K 上的连续函数空间

$\mathcal{C}[0,\omega_1]$　$[0,\omega_1]$ 上的连续函数，其中 ω_1 是第一不可数序数

$\mathcal{C}_0$　具有紧支集的连续函数空间

$\mathcal{C}^k, 1 \leqslant k \leqslant \infty$,　具有 k 次连续导数的可微函数空间

$\mathcal{C}^{1.1}$ 具有 *Lipschitz* 次连续导数的可微函数空间

$\mathcal{L}^p([a,b];X), 1 \leqslant p \leqslant \infty$， X- 值映射的标准 Lebesgue 空间

$W^{1,p}, H^p$ 标准 Sobolev 空间

$\mathcal{M}, \mathcal{M}_b$ 测度空间 (连续函数空间的对偶)

BV 有界变差函数空间

c 具有上确界范数的实数序列的空间

c_0 所有收敛于 0 的序列组成的 c 的子空间

$l^p, 1 \leqslant p \leqslant \infty$， 具有 $p-$ 范数的实数序列空间

集合

$\varnothing$ 空集

$\mathbb{N}$ 自然数集

$B_r(x)$ 球心在 x 半径为 r 的球

$\mathbb{B}_X$ 空间 X 的闭单位球

$\mathbb{B}, \mathbb{B}^*$ 空间及其对偶的闭单位球

S, S^* 空间及其对偶的单位球面

$\mathrm{int}\,\Omega, \mathrm{ri}\,\Omega$ Ω 的内部和相对内部

$\mathrm{cl}\,\Omega, \mathrm{cl}^*\,\Omega$ Ω 的闭包和弱 * 拓扑闭包

$\mathrm{bd}, \partial\Omega$ 集合的边界

$\mathrm{co}\Omega, \mathrm{clco}\Omega$ 凸包和闭凸包

$\mathrm{cone}\Omega$ 锥包

$\mathrm{aff}\Omega, \overline{\mathrm{aff}\Omega}$ 仿射包和闭仿射包

$\mathrm{mes}\Omega, \mathcal{L}^n(\Omega)$ Lebesgue(n-维) 测度

$\Pi(x;\Omega)$ x 在 Ω 上的投影

$T(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的相依切锥

$T_W(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的弱相依切锥

$T_C(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的 Clarke 切锥

$N(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的基本/极限法锥

$N_+(\bar{x};\Omega(\bar{y}))$ $\Omega(\bar{y})$ 在 $\bar{x}$ 的增广极限法锥

$\widehat{N}(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的预法锥或 Fréchet 法锥

$N_C(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的 Clarke 法锥

$N_G(\bar{x};\Omega), \tilde{N}_G(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的近似 G- 法锥及其核

$N_P(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的迫近法锥

$\widehat{N}_\varepsilon(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的 ε- 法向量的集合

$S_\varepsilon(\bar{x};\Omega)$ Ω 在 $\bar{x}$ 的 ε- 支撑

函数

$\delta(\cdot;\Omega)$ 集合的指示函数

$\text{dist}(\cdot;\Omega)$, $d_\Omega(\cdot)$ 距离函数

$\rho(x,y):=\text{dist}(y;F(x))$ 增广距离函数

$\text{dom}\,\varphi$ $\varphi:X\to\overline{\mathbb{R}}$ 的有效域

$\text{epi}\varphi$, $\text{hypo}\varphi$, $\text{gph}\varphi$ φ 的上图、下图和图

$x\xrightarrow{\varphi}\bar{x}$ $x\to\bar{x}$ 且 $\varphi(x)\to\varphi(\bar{x})$

$\mathcal{H}$ 最优控制中的 Hamilton 函数

H 最优控制中的 Hamilton-Pontryagin 函数

L 最优化中的 Lagrange 函数

L_Ω 相对于 Ω 的基本 Lagrange 函数

$\tau(F;h)$ 连续性的平均模/连续性均模

$\varphi'(\bar{x})$, $\nabla\varphi(\bar{x})$ φ 在 $\bar{x}$ 的 Fréchet 导数/梯度

$\varphi'_\beta(\bar{x})$, $\nabla_\beta\varphi(\bar{x})$ φ 在 $\bar{x}$ 相对于某生成族的导数/梯度

$|\nabla\varphi|(\bar{x})$ φ 在 $\bar{x}$ 的 (强) 斜率

$\varphi'(\bar{x};v)$ φ 在 $\bar{x}$ 沿方向 v 的经典方向导数

$\varphi^\circ(\bar{x};v)$, $\varphi^\uparrow(\bar{x};v)$ φ 的广义方向导数和次导数

$d^-\varphi(\bar{x};v)$, $d^+\varphi(\bar{x};v)$ φ 的 Dini-Hadamard 下/上方向导数

$\partial\varphi(\bar{x})$ φ 在 $\bar{x}$ 的基本/极限次微分

$\partial^+\varphi(\bar{x})$ φ 在 $\bar{x}$ 的上次微分

$\partial^0\varphi(\bar{x})$ φ 在 $\bar{x}$ 的对称次微分

$\partial_\geqslant\varphi(\bar{x})$ φ 在 $\bar{x}$ 的右边次微分

$\partial^\infty\varphi(\bar{x})$ φ 在 $\bar{x}$ 的奇异次微分

$\widehat{\partial}\varphi(\bar{x})$, $\widehat{\partial}^+\varphi(\bar{x})$ φ 在 $\bar{x}$ 的 Fréchet 次微分和上次微分

$\partial_A\varphi(\bar{x})$, $\nabla_G\varphi(\bar{x})$ φ 在 $\bar{x}$ 的近似 A- 次微分和 G- 次微分

$\partial_C\varphi(\bar{x})$ φ 在 $\bar{x}$ 的 Clarke 次微分/广义梯度

$\partial_\beta\varphi(\bar{x})$ φ 在 $\bar{x}$ 的黏性 (生成族的)β- 次微分

$\partial_P\varphi(\bar{x})$ φ 在 $\bar{x}$ 的邻近次微分

$\widehat{\partial}_\varepsilon\varphi(\bar{x})$, $\widehat{\partial}_{a\varepsilon}\varphi(\bar{x})$, $\widehat{\partial}_{g\varepsilon}\varphi(\bar{x})$ φ 在 $\bar{x}$ 的 Fréchet 型 ε- 次微分

$\partial^-_\varepsilon\varphi(\bar{x})$ φ 在 $\bar{x}$ 的 Dini ε- 次微分

$\nabla^2\varphi(\bar{x})$ φ 在 $\bar{x}$ 的经典 Hesse 阵 (在 $\mathbb{R}^n$ 中则为二阶导数矩阵)

$\partial^2\varphi$, $\partial^2_N\varphi$, $\partial^2_M\varphi$ φ 的二阶次微分 (广义 Hesse 阵)

映射

$f:X\to Y$ 从 X 到 Y 的单值映射

$F:X\rightrightarrows Y$ 从 X 到 Y 的集值映射

$\text{dom}\,F$ F 的有效域

$\text{rge}\,F$ F 的值域

$\text{gph}\,F$ F 的图

$\ker F$ F 的核

$F^{-1}: Y \rightrightarrows X$ $F: X \rightrightarrows Y$ 的逆映射

$F(\Omega)$, $F^{-1}(\Omega)$ Ω 在 F 下的像和逆像/预像

$F \circ G$ 映射的复合

$F h_\circ G$ 映射的 h-复合

$\Delta(\cdot;\Omega)$ 集合的指示映射

Ω_ρ 集合的增大映射

E_φ 上图映射

$\varepsilon(f,\Theta)$ $f: X \to Y$ 相对于 $\Theta \subset Y$ 的广义上图

$DF(\bar{x},\bar{y})$ F 在 $(\bar{x},\bar{y}) \in \operatorname{gph} F$ 的图/相依锥导数

$D^*F(\bar{x},\bar{y})$ F 在 $(\bar{x},\bar{y}) \in \operatorname{gph} F$ 的 (基本) 上导数

$D_N^*F(\bar{x},\bar{y})$ F 在 $(\bar{x},\bar{y})$ 的基本上导数

$D_M^*F(\bar{x},\bar{y})$, $\widetilde{D}_M^*F(\bar{x},\bar{y})$ F 在 $(\bar{x},\bar{y})$ 的混合和逆混合上导数

$\widehat{D}^*F(\bar{x},\bar{y})$, $\widehat{D}_\varepsilon^*F(\bar{x},\bar{y})$ F 在 $(\bar{x},\bar{y})$ 的 Fréchet 上导数和 $\varepsilon-$ 上导数

$Jf(\bar{x})$ f 在 $\bar{x}$ 的广义 Jacob 矩阵

$\Lambda f(\bar{x})$ f 在 $\bar{x}$ 的导容

索　　引

《现代数学译丛》已出版书目

(按出版时间排序)

1 椭圆曲线及其在密码学中的应用——导引 2007.12 〔德〕Andreas Enge 著 吴 铤 董军武 王明强 译
2 金融数学引论——从风险管理到期权定价 2008.1 〔美〕Steven Roman 著 邓欣雨 译
3 现代非参数统计 2008.5 〔美〕Larry Wasserman 著 吴喜之 译
4 最优化问题的扰动分析 2008.6 〔法〕J. Frédéric Bonnans 〔美〕Alexander Shapiro 著 张立卫 译
5 统计学完全教程 2008.6 〔美〕Larry Wasserman 著 张 波 等译
6 应用偏微分方程 2008.7 〔英〕John Ockendon, Sam Howison, Andrew Lacey & Alexander Movchan 著 谭永基 程 晋 蔡志杰 译
7 有向图的理论、算法及其应用 2009.1 〔丹〕J. 邦詹森 〔英〕G. 古廷 著 姚兵 张忠辅 译
8 微分方程的对称与积分方法 2009.1〔加〕乔治 W. 布卢曼 斯蒂芬 C. 安科 著 闫振亚 译
9 动力系统入门教程及最新发展概述 2009.8 〔美〕Boris Hasselblatt & Anatole Katok 著 朱玉峻 郑宏文 张金莲 阎欣华 译 胡虎翼 校
10 调和分析基础教程 2009.10 〔德〕Anton Deitmar 著 丁勇 译
11 应用分支理论基础 2009.12 〔俄〕尤里・阿・库兹涅佐夫 著 金成桴 译
12 多尺度计算方法——均匀化及平均化 2010.6 Grigorios A. Pavliotis, Andrew M. Stuart 著 郑健龙 李友云 钱国平 译
13 最优可靠性设计：基础与应用 2011.3 〔美〕Way Kuo, V. Rajendra Prasad, Frank A.Tillman, Ching-Lai Hwang 著 郭进利 闫春宁 译 史定华 校
14 非线性最优化基础 2011.4 〔日〕Masao Fukushima 著 林贵华 译
15 图像处理与分析：变分，PDE，小波及随机方法 2011.6 Tony F. Chan, Jianhong (Jackie) Shen 著 陈文斌，程晋 译
16 马氏过程 2011.6 〔日〕福岛正俊 竹田雅好 著 何萍 译 应坚刚 校

17 合作博弈理论模型 2011.7 〔罗〕Rodica Branzei 〔德〕Dinko Dimitrov 〔荷〕Stef Tijs 著 刘小冬 刘九强 译

18 变分分析与广义微分 I：基础理论 2011. 9 〔美〕 Boris S. Mordukhovich 著 赵亚莉 王炳武 钱伟懿 译

19 随机微分方程导论应用(第 6 版) 2012.4 〔挪〕Bernt Øksendal 著 刘金山 吴付科 译

20 金融衍生产品的数学模型 2012.4 郭宇权(Yue-Kuen Kwok) 著 张寄洲 边保军 徐承龙 等 译

21 欧拉图与相关专题 2012.4 〔英〕Herbert Fleischner 著 孙志人 李 皓 刘桂真 刘振宏 束金龙 译 张 昭 黄晓晖 审校

22 重分形：理论及应用 2012.5 〔美〕戴维·哈特 著 华南理工分形课题组 译

23 组合最优化：理论与算法 2014. 1 〔德〕 Bernhard Korte Jens Vygen 著 姚恩瑜 林治勋 越民义 张国川 译

24 变分分析与广义微分 II：应用 2014. 1 〔美〕 Boris S. Mordukhovich 著 李 春 王炳武 赵亚莉 王 东 译

(O-5350.01)

科学数理分社
电话：(010)64033664
E-mail：math-phy@mail.sciencep.com
网址：http://www.math-phy.cn

销售分类建议：高等数学

www.sciencep.com



定 价：138.00 元